(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 February 2003 (06.02.2003)

PCT

(10) International Publication Number
WO 03/010308 A2

(51) International Patent Classification[7]: C12N 15/00

(21) International Application Number: PCT/GB02/03336

(22) International Filing Date: 19 July 2002 (19.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0117964.7 24 July 2001 (24.07.2001) GB

(71) Applicant *(for all designated States except US)*: GENE EXPRESSION TECHNOLOGIES LIMITED [GB/GB]; 90 Fetter Lane, London EC4A 1JP (GB).

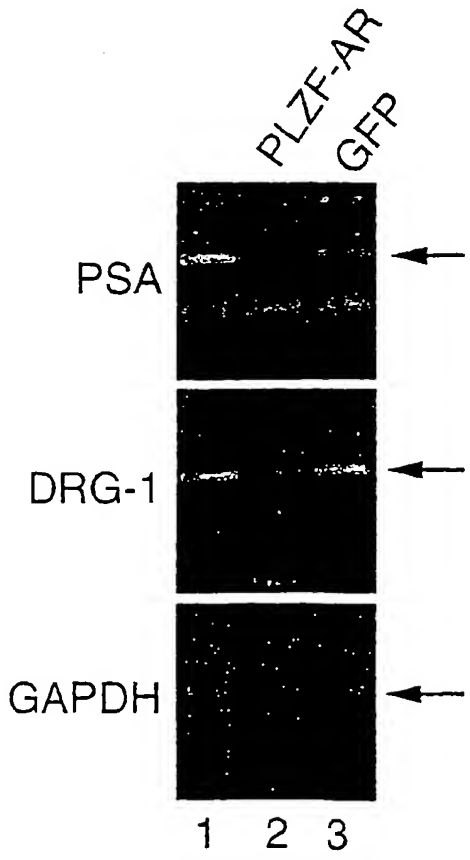
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **BULUWELA, Lakjaya** [GB/GB]; Imperial College School of Medicine, Dept. of Cancer Medicine, Du Cane Road, London W12 0NN (GB). **HOLMES, David** [GB/GB]; Imperial College School of Medicine, Dept. of Cancer Medicine, Du Cane Road, London W12 0NN (GB). **KAMALATI, Tahereh** [GB/GB]; Imperial College School of Medicine Dept. of Cancer Medicine, Du Cane Road, London W12 0NN (GB). **WAXMAN, Jonathan** [GB/GB]; Imperial College School of Medicine, Dept. of Cancer Medicine, Du Cane Road, London W12 0NN (GB). **ALI, Simak** [GB/GB]; Imperial College School of Medicine, Dept. of Cancer Medicine, Du Cane Road, London W12 0NN (GB).

(74) Agent: **MILES, John, S.**; Eric Potter Clarkson, Park View House, 58 The Ropewalk, Nottingham NG1 5DD (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,

[Continued on next page]

(54) Title: CONTROL OF GENE EXPRESSION



(57) Abstract: A method of suppressing the expression of a selected endogenous gene in a eukaryotic cell the method comprising introducing into the cell (a) a polypeptide comprising a nucleic acid binding portion which binds to a site at or associated with the selected gene and a modifying portion which comprises a polypeptide or peptidomimetic which is capable of modulating covalent modification of nucleic acid or chromatin, for example a chromatin inactivation portion, or (b) a polynucleotide encoding said polypeptide. The binding of the molecule to nucleic acid may be modulated by a ligand and the molecule may comprise a ligand binding portion. The nucleic acid binding portion may be a nucleic acid binding portion of a nuclear receptor DNA binding protein.

WO 03/010308 A2

GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

CONTROL OF GENE EXPRESSION

The present invention relates to the control of gene expression and, in particular, it relates to methods of, and means for, suppressing the expression of a particular, selected gene.

The ability to selectively suppress the expression of a gene is useful in many areas of biology, for example in methods of treatment where the expression of the gene may be undesirable; in preparing models of disease where lack of expression of a particular gene is associated with the disease; in modifying the phenotype in order to produce desirable properties. Thus, the ability selectively to suppress the expression of a gene may allow the "knockout" of human genes in human cells (whether wild type or mutant) and the knockout of eukaryotic genes in studies of development and differentiation.

Present methods of attempting to suppress the expression of a particular gene fall into three main categories, namely antisense technology, ribozyme technology and targeted gene deletion brought about by homologous recombination.

Antisense techniques rely on the introduction of a nucleic acid molecule into a cell which typically is complementary to a mRNA expressed by the selected gene. The antisense molecule typically suppresses translation of the mRNA molecule and prevents the expression of the polypeptide encoded by the gene. Modifications of the antisense technique may prevent the transcription of the selected gene by the antisense molecule binding to the gene's DNA to form a triple helix.

Ribozyme techniques rely on the introduction of a nucleic acid molecule into a cell which expresses a RNA molecule which binds to, and catalyses the selective cleavage of, a target RNA molecule. The target RNA molecule is typically a mRNA molecule, but it may be, for example, a retroviral RNA molecule.

Antisense- and ribozyme-based techniques have proven difficult to implement and they show varying degrees of success in target gene suppression or inactivation. Furthermore, these two techniques require persistent expression or administration of the gene-inactivating agent.

Targeted gene deletion by homologous recombination requires two gene-inactivating events (one for each allele) and is not easily applicable to primary cells, particularly for example primary human mammary cells which can only be maintained in culture for a few passages. Targeted gene deletion has remained difficult to perform in plants. The *cre-lox* mediated site-specific integration has been the method of choice although the efficiency of specific integrative events is low (Alberts *et al* (1995) *Plant J.* **7**, 649-659; Vergunst & Hooykass (1998) *Plant Mol. Biol.* **38**, 393-406; Vergunst *et al* (1998) *Nucl. Acids Res.* **26**, 2729-2734).

These major shortcomings in existing technology have led us to seek an alternative strategy.

Acute promyelocytic leukaemia (APL) is underlined by the involvement of mutant retinoic acid receptor (RAR) proteins, formed by gene fusions brought about by chromosomal translocations. Molecular analysis of one APL subset has identified a fusion between the RARI gene and a Kruppel-like zinc finger gene named promyelocytic leukaemia zinc finger (PLZF).

Further investigations have shown that the resulting PLZF-RARI fusion protein probably functions as a gene repressor by targeting histone deacetylation of retinoic acid regulated genes. Several studies have shown that this repression is mediated by the PLZF portion of the fusion protein, which interacts with a complex of proteins which includes the components N-CoR, SMRT, Sin3 and HDAC and which in turn results in the recruitment of the histone deacetylase (HDAC) complex to target genes (see, for example, Grignani *et al* (1998) *Nature* **391**, 815-818; Chen *et al* (1993) *EMBO J* **12**, 1161-1167; Razin (1998) *EMBO J.* **17**, 4905-4908; David *et al* (1998) *Oncogene* **16**, 2549-2556; and Lin *et al* (1998) *Nature* **391**, 811-814). HDAC directed gene inactivation, therefore results from the targeted assembly of components, some of which have been identified (eg N-CoR, SMRT, Sin3 etc) making a gene inactivating complex which mediates its effect through histone deacetylation.

Although this work shows that in certain forms of APL fusion proteins are able to recruit histone deacetylase activity which appears to have the effect of inactivating the expression of certain genes, no-one has demonstrated selective suppression of expression of a chosen target endogenous gene or a set of endogenous genes using a method based on recruitment of histone deacetylase or other means of inactivating chromatin. We have shown that this can be achieved.

Beerli *et al* (1998) *PNAS* **95**, 14628-14633 report repression of a transiently expressed luciferase reporter gene construct using a zinc finger protein fused to either a KRAB (Krüppel-associated box), ERD (ERF repressor domain) or SID (mSIN3 interaction domain) transcriptional repressor domain.

WO01/02019 reports repression of a reporter gene using a PLZF-estrogen receptor or PLZF-androgen receptor fusion polypeptide.

A first aspect of the invention provides a method of suppressing the expression of a selected endogenous gene in a eukaryotic cell, the method comprising introducing into the cell (a) a molecule comprising a nucleic acid binding portion which binds to a site at or associated with the selected endogenous gene and a modifying portion which comprises a polypeptide or peptidomimetic which is capable of modulating covalent modification of nucleic acid or chromatin, or (b) a polynucleotide encoding said molecule.

It is preferred that the modifying portion is a chromatin inactivation portion.

Alternatively, the modifying portion may be a portion that is capable of modulating covalent modification, for example methylation, of nucleic acid, preferably DNA. Thus, the modifying portion may be or comprise a DNA modifying enzyme, or may be capable of recruiting such an enzyme. The modulation has the effect of suppressing the selected gene.

The modifying portion does not change the sequence of the nucleic acid. It is preferred that the modifying portion does not cleave the nucleic acid backbone. The modifying portion is preferably not a recombinase or a restriction endonuclease.

For example, the modifying portion may comprise (or be capable of recruiting) a methyl transferase or a component of a methyltransferase complex, for example as discussed in Okano M, Xie S, Li E. (1998) Cloning and characterization of a family of novel mammalian DNA

(cytosine-5) methyltransferases. *Nat Genet* 19:219-220; Adrian P. Bird and Alan P. Wolffe (1999) Methylation-Induced Repressionó Belts, Braces, and Chromatin. *Cell* 99, 451-454. The modifying portion may comprise a methylase that methylates CpG dinucleotides, for example the methylase M.*Sss*I, a methyltransferase from *Spiroplasma* species (see Xu & Bestor (1997) *Nature Genet* 27, 304-308) The modifying portion may comprise a DNA adenine methyltransferase (van Steensel *et al* (2001) *Nature Genet* 27, 304-308), though this is not preferred.

As will be appreciated, the site is one which is naturally present in a eukaryotic genome.

Specific regulation of an endogenous gene in its native chromatin environment using a modifying portion as defined above, for example a chromatin inactivation portion, has not previously been demonstrated.

It is preferred that the molecule is a polypeptide or comprises a polypeptide portion; for example it is strongly preferred that the modifying portion, for example chromatin inactivation portion, is a polypeptide. It is further preferred that the molecule or polypeptide is a hybrid molecule or polypeptide which does not occur in nature. For example, it is preferred if the nucleic acid binding portion is derived from one protein and the modifying or chromatin inactivation portion is derived from a different protein and that the molecular configuration does not arise in nature, for example through chromosome translocation events. The proteins from which the nucleic acid binding portion and the modifying or chromatin inactivation portion are derived may be from the same species (for example, as is described in more detail below, the nucleic acid binding portion may be a DNA binding portion of a human steroid

receptor protein such as oestrogen receptor (ER) and the chromatin inactivation portion may be a portion of human PLZF) or they may be from different species (for example a bacterial DNA binding protein may be fused to a portion of human PLZF).

Thus, in a particular preferred embodiment the polypeptide is one which is produced by genetic engineering means wherein the nucleic acid binding portion and the modifying or chromatin inactivation portion are selected as is described in more detail below.

It is preferred if the nucleic acid binding portion is not the *Saccharomyces cerevisiae* GAL4 protein or a DNA-binding portion thereof, and it is preferred if the nucleic acid binding portion is not the *Escherichia coli* LexA protein or a DNA-binding portion thereof.

The term "endogenous gene" refers to a gene that is native to the cell ie which is not heterologous to the cell and is in its natural genomic context. The site present in a eukaryotic genome is a site which is at or associated with a selected gene or genes whose expression it is desirable to suppress or inactivate. The site is a site which is naturally present in a eukaryotic genome and is in its natural genomic context.

By "genome" we include not only nuclear chromosomal DNA but other DNA naturally present in the eukaryotic cell, such as plastid DNA.

It is preferred if the nucleic acid binding portion can bind to chromosomal DNA or, as is described in more detail below, to RNA transcribed from chromosomal DNA.

The chromatin inactivation portion may be any polypeptide or part thereof which directly or indirectly leads to chromatin inactivation. By "directly" leading to chromatin inactivation we mean that the polypeptide or part thereof itself acts on the chromatin to inactivate it. By "indirectly" leading to chromatin inactivation we mean that the polypeptide or part thereof does not itself act on the chromatin but rather it is able to recruit or promote a cellular component to do so. Equivalent preferences apply in relation to other modifying portions.

Chromatin inactivation generally results in the suppression or inactivation of gene expression. Chromatin inactivation is typically a localised event such that suppression or inactivation of gene expression is restricted to, typically, one or a few genes. Thus, the chromatin inactivation portion is any suitable polypeptide which, when part of the polypeptide of the invention and when targeted to a selected gene by the nucleic acid binding portion, locally inactivates the chromatin associated with the selected gene so that expression of the gene is inactivated or suppressed. Histone deacetylation is associated with chromatin inactivation and so it is particularly preferred if the chromatin inactivation portion facilitates histone deacetylation. Targeted deacetylation of histones associated with a given gene leads to gene inactivation in an, essentially, irreversible manner. By "suppression" or "inactivation" of gene expression we mean that in the presence of the molecule or polypeptide the expression of the selected, targeted gene is at least five-fold, preferably at least ten-fold, more preferably at least 100-fold, and most preferably at least 1000-fold lower than in the absence of the molecule or polypeptide of the under equivalent conditions. Gene expression can be measured using any suitable method including using reverse transcriptase-polymerase chain reaction (RT-PCR), RNA hybridisation, RNAse protection assays, nuclear

run-off assays and alteration of chromatin as judged by DNAse 1 hypersensitivity.

In animal and plant cells histone deacetylation is brought about by the so-called histone deacetylase complex (HDAC) which contains, in addition to one or more histone deacetylase enzymes, ancillary proteins which are involved in the formation and function of the complex. In addition, there are other protein components which although they may not be part of HDAC they bind to or otherwise interact with HDAC and help facilitate histone deacetylation.

Deacetylation and acetylation of histones is a well-known phenomenon which is reviewed in the following: Chen & Li (1998) *Crit. Rev. Eukaryotic Gene Expression* **8**, 169-190; Workman & Kingston (1998) *Ann. Rev. Biochem.* **67**, 545-579; Perlmann & Vennstrom (1995) *Nature* **377**, 387- ; Wolfe (1997) *Nature* **387**, 16-17; Grunstein (1997) *Nature* **389**, 349-352; Pazin & Kadonaga (1997) *Cell* **89**, 325-328; DePinho (1998) *Nature* **391**, 533-536; Bestor (1998) *Nature* **393**, 311-312; and Grunstein (1998) *Cell* **93**, 325-328.

The polypeptide composition of the HDAC complex is currently under investigation. Polypeptides which may form part of, or are associated with, certain HDAC complexes include histone deacetylase 1 (HDAC1) Taunton *et al* (1996) *Nature* **272**, 408-441); histone deacetylase 2 (HDAC2) (Yang *et al* (1996) *Proc. Natl. Acad. Sci. USA* **93**, 12845-12850); histone deacetylase 3 (HDAC3) (Dangond *et al* (1998) *Biochem. Biophys. Res. Comm.* **242**, 648-652); N-CoR (Horlein *et al* (1995) *Nature* **377**, 397-404); SMRT (Chen & Evans (1995) *Nature* **377**, 454-457); SAP30 (Zhang *et al* (1998) *Molecular Cell* **1**, 1021-1031). Sin3 (Ayer *et*

al (1995) *Cell* **80**, 767-776; Schreiber-Agus *et al* (1995) *Cell* **80**, 777-786) SAP18 (Zhang *et al* (1997) *Cell* **89**, 357-364); and RbAp48 (Qian *et al* (1993) *Nature* **364**, 648-652). All of these papers are incorporated herein by reference. It is believed that there may be further components of the HDAC complex or which interact with the HDAC complex which are, as yet, undiscovered. It is envisaged that these too will be useful in the practice of the invention.

PLZF has been shown to interact with N-CoR and SMRT, which in turn recruit a HDAC complex. PLZF will also directly interact with HDAC (Lin *et al* (1998) *Nature* **391**, 811-814; Grignani *et al* (1998) *Nature* **391**, 815-818; David *et al* (1998) *Oncogene* **16**, 2549-2556).

Complexes formed which contain any of N-CoR, SMRT, Sin3, SAP18, SAP30 and histone deacetylase are described in Heinzel *et al* (1997) *Nature* **387**, 43-48; Alland *et al* (1997) *Nature* **387**, 49-55; Hassig *et al* (1997) *Cell* **89**, 341-347; Laherty *et al* (1997) *Cell* **89**, 349-356; Zhang *et al* (1997) *Cell* **89**, 357-364; Kadosh & Struhl (1997) *Cell* **89**, 365-371; Nagy *et al* (1997) *Cell* **89**, 373-380; and Laherty *et al* (1998) *Molecular Cell* **2**, 33-42. All of these papers are incorporated herein by reference.

Thus, it is particularly preferred if the component of a HDAC complex or the polypeptide which binds to or facilitates recruitment of a HDAC complex is any one of PLZF, SMRT, Sin3, SAP18, SAP30 or N-CoR, or HDACs including HDAC1, HDAC2 or HDAC3. It will be appreciated that it may not be necessary for all of the polypeptides to be present so long as a functional portion thereof is present. For example, with respect to histone deacetylase enzymes (for example, HDAC1, HDAC2 or HDAC3) the functional portion may be a portion that retains histone

deacetylase activity or it may be a portion which contains a binding site for other components of a HDAC complex or a portion which otherwise recruits the HDAC complex and promotes histone deacetylation. Similarly, with respect to other components of the HDAC complex or polypeptides which bind to the HDAC complex the functional portion may be a portion which contains a binding site for other components of the HDAC complex. To date six mammalian HDAC genes have been described (Grozinger *et al* (1999) *Proc. Natl. Acad. Sci. USA* **96**, 4868-4873), it is believed that any one or more of these may be useful in the practise of the present invention.

It is preferred that the chromatin inactivation portion is not N-CoR or a portion thereof; or the C-terminal domain of the vErbA, T_3R (including $T_3R\beta1$ or $T_3R\alpha$) or RAR (including $RAR\alpha$) receptor molecule, particularly if the nucleic acid binding portion is the *Saccharomyces cerevisiae* GAL4 protein or a DNA-binding portion thereof, or the *Escherichia coli* LexA protein or a DNA-binding portion thereof. It is preferred that when the nucleic acid binding portion is a DNA binding portion of RARI the chromatin inactivation portion is not a portion of PLZF protein and is not a portion of PML protein; and that when the nucleic acid binding portion is a DNA binding portion of the *Saccharomyces cerevisiae* GAL4 protein the chromatin inactivation portion is not a portion of PLZF protein, the C-terminal domain of vErbA, T_3R, $T_3R\beta1$ or RAR, or N-CoR or a portion of N-CoR, or CIR (CBF1 Interacting coRepressor; see Hsieh *et al* (1999) *PNAS* **96**, 23-28); and that when the nucleic acid binding portion is a DNA binding portion of the *Escherichia coli* LexA the chromatin inactivation portion is not mSin3, or the C-terminal domain of $T_3R\alpha$ or $RAR\alpha$.

For the avoidance of doubt, KRAB is not included within the meaning of the term "modifying portion" or "chromatin inactivation portion". KRAB is a transcriptional repressor whose mode of action is considered to involve mechanisms other than chromatin inactivation. Although not preferred, any fragment of KRAB that, when part of the molecule/polypeptide as defined above and when targeted to a selected gene by the nucleic acid binding portion, locally inactivates the chromatin associated with the selected gene so that expression of the gene is inactivated or suppressed, is included within the term "chromatin inactivation portion". For example, any fragment of KRAB that is capable of binding to or facilitating recruitment of a HDAC complex is included within the term "chromatin inactivation portion". However, any such fragments are not preferred.

It is believed that binding motifs are present within the components of the HDAC complex or within polypeptides which bind the HDAC complex and these motifs may be sufficient to act as chromatin inactivation portions in the polypeptide of the invention since they may facilitate histone deacetylation by recruiting a HDAC complex.

Furthermore, it will be appreciated that variants of a component of the HDAC complex or variants of a polypeptide which binds to the HDAC complex may be used. Suitable variants include not only functional portions as described above, but also variants in which amino acid residues have been deleted or replaced or inserted provided that the variant is still able to facilitate histone deacetylation. Thus, suitable variants include variants of histone deacetylase in which the amino acid sequence has been modified compared to wild-type but which retain their ability to deacetylate histones. Similarly, suitable variants include variants

of, for example, Sin3 or PLZF in which the amino acid sequence has been modified compared to wild-type but which retain their ability to interact with or in the HDAC complex. Similarly, variants of other proteins interacting with components of the HDAC complex and other transcription factors that can bring about gene inactivation through HDAC activity may be used.

All or parts of the Rb, MAD and MeCpG2 proteins may interact with the HDAC complex.

While most work has been done on HDAC complexes and polypeptides involved in recruiting HDAC complexes in mammalian systems, the fundamental nature of the system is such that functionally equivalent polypeptides are expected to be found in other eukaryotic cells, in particular in other animal cells and in plant cells. For example, Figure 5 of WO01/02019 shows that polypeptides very closely related to human HDAC1 are present in arabidopsis and in yeast. A plant HDAC complex has been isolated from maize (Lussen *et al* (1997) *Science* **277**, 88-91) and a comparative study of histone deacetylases from plant, fungal and vertebrate cells has been undertaken (Lechner *et al* (1996) *Biochim. Biophys. Acta* **1296**, 181-188). Histone deacetylase inhibitors have been shown to derepress silent rRNA genes in Brassica (Chen & Pickard (1997) *Genes Dev.* **11**, 2124-2136) and a naturally occurring host selective toxin (HC toxin) from *Cochliobolus carbonum* inhibits plant, fungal and mammalian histone deacetylases (Brosch *et al* (1995) *Plant Cell* 7, 1941-1950).

It is not necessary that the modifying or chromatin inactivation portion is from the same cell type or species as the cell into which the molecule or

polypeptide (or polynucleotide encoding the polypeptide) is introduced but it is desirable if it is since such a modifying or chromatin inactivation portion may be able to inactivate nucleic acid or chromatin more effectively in that cell.

It is particularly preferred if the chromatin inactivation portion of the polypeptide is PLZF, a portion of PLZF that can facilitate histone deacetylation, or a polypeptide, or portion of a polypeptide, known to cause gene activation *via* histone deacetylation. For example, the portion of PLZF in PLZF-RARI which is involved in APL is believed to interact with N-CoR and SMRT.

It is also particularly preferred if the chromatin inactivation portion is a polypeptide with histone deacetylase enzyme activity such as contained in HDAC1, HDAC2 or HDAC3.

The nucleic acid binding portion may be any suitable binding portion which binds to a site present in a eukaryote, such as a plant or animal, genome, which site is at or associated with the selected gene whose expression is to be suppressed by the presence of the chromatin inactivating portion of the molecule or polypeptide. It is preferred that the nucleic acid binding portion binds selectively to the desired site. There may be one or more desired sites to which the nucleic acid binding portion may bind. For example, the molecule or polypeptide may be used to suppress the expression of a group of genes which each have a binding site for a common DNA binding portion (for example, are under the controls of a steroid hormone receptor such as the oestrogen receptor (ER) or androgen receptor (AR)). As noted above, the site to which the nucleic

acid binding portion binds is naturally present in the eukaryotic cell and is present in its natural position in the genome.

The nucleic acid binding portion may be a DNA binding portion or an RNA binding portion. Molecules, for example proteins, which have the ability to bind either DNA or RNA in a sequence selective manner are well known in the art and some are described in more detail below. In the case of the RNA binding portion, the site present in the eukaryotic genome which binds the RNA binding portion is, typically, nascent RNA being transcribed from DNA at the selected site for inactivation. The RNA may be that which is being transcribed by the gene whose expression is to be suppressed, or it may be that which is being transcribed by a gene adjacent to, or at least close to, the gene whose expression is to be suppressed. It is preferred that the RNA binding portion binds to an RNA sequence which is at or close to the 5N end of the transcript. It will be appreciated that whilst being transcribed, nascent RNA remains at or close to its site of transcription and that if the site of transcription is at or close to the gene whose expression is to be suppressed, using an RNA binding portion in the polypeptide of the invention facilitates the localisation of the chromatin inactivation portion to the desired site.

The DNA binding portion may be all or a DNA-binding portion of a zinc-finger DNA binding protein or it may be all or a DNA-binding portion of a helix-turn-helix DNA binding protein.

Suitably the DNA binding portion may be all or a DNA-binding portion of an animal or plant DNA binding protein, or it may be all or a DNA binding portion of a bacterial or yeast DNA binding protein which has been engineered to bind to one or more sites in the plant or animal

genome. Bacterial or yeast DNA binding proteins are less preferred and it is particularly preferred if the DNA binding protein does not contain a DNA binding portion of wild-type *Saccharomyces cerevisiae* GAL4 or wild-type *Escherichia coli* LexA.

Any DNA binding protein with the ability to bind DNA through a DNA recognition sequence may be used. This includes DNA binding proteins, and engineered DNA binding proteins, such as engineered zinc finger proteins and helix-turn-helix DNA binding proteins.

Databases listing transcription factors and their binding sites are listed below:

http://www.embl-heidelberg.de/srs5bin/cgi-bin/wgetz?-fun+pagelibinfo+-info+TFFACTOR
http://www.embl-heidelberg.de/srs5bin/cgi-bin/wgetz?-fun+pagelibinfo+-info+TFSITE
http://www.embl-heidelberg.de/srs5bin/cgi-bin/wgetz?-fun+pagelibinfo+-info+TFCELL
http://www.embl-heidelberg.de/srs5bin/cgi-bin/wgetz?-fun+pagelibinof+-info+TFCLASS
http://www.embl-heidelberg.de/srs5bin/cgi-bin/wgetz?-fun+pagelibinfo+-info+TFMATRIX
http://www.embl-heidelberg.de/srs5bin/cgi-bin/wgetz?-fun+pagelibinfo+-info+TFGENE

It is believed that all or part of the listed transcription factors may be useful in the practice of the invention.

Other gene regulatory proteins which may be useful in the practice of the invention include virally encoded DNA binding proteins (or variants thereof) such as those required for regulating viral and cellular gene expression and/or viral replication. These include but are not limited to the large T antigen of polyoma viruses, the E2 protein of papillomaviruses and the ICP4, ICP0 protein of herpesviruses.

Sequence specific RNA binding proteins, which bind to nascent RNA, may be engineered to bring about gene inactivation by the methods of the invention through HDAC complex formation in the proximity of transcriptionally active chromatin. For example, the transcriptionally active chromatin may be proviral (though it is preferred that the chromatin is endogenous to the cell and is not proviral) and the RNA binding protein one which binds to transcribed proviral RNA. The *tat* protein of HIV is an example of an RNA binding protein.

In plants, DNA binding proteins are involved in, amongst other things, floral development, cold regulation/adaptation, and plant responses to ethylene or pathogens. Thus, the methods of the invention may be used to analyse the role of these genes in these developmental and other processes.

A particularly preferred embodiment is wherein the DNA binding portion is all or a DNA binding part of a nuclear receptor DNA binding protein such as a steroid hormone receptor protein. Features of a nuclear receptor DNA binding protein will be well known to those skilled in the art, for example characteristic domains.

In a further preferred embodiment the DNA binding portion is all or a DNA binding part of a nuclear receptor DNA binding protein such as a

steroid hormone receptor protein and the modifying portion is a chromatin inactivating portion.

In a still further preferred embodiment the DNA binding portion is all or a DNA binding part of a nuclear receptor DNA binding protein such as a steroid hormone receptor protein and the modifying portion is a DNA modifying portion, for example comprises (or is capable of recruiting) all or a part of a methyltransferase or methyltransferase complex component.

The nuclear receptor DNA binding protein superfamily includes oestrogen receptor (ER), androgen receptor (AR), progesterone receptor (PR), retinoic acid receptor (RAR) and the like (see Mangelsdorf *et al* (1995) *Cell* 83, 835-839 for a review and nomenclature).

Databases listing putative nuclear receptors are listed below:

http://www.toulouse.inra.fr/prodom/cgi-bin/ReqProdomII.pl?id_dom2=PD00
0035&prodom_release=2001.1

http://ca.expasy.org/cgi-bin/nicesite.pl?PS00031

Nuclear Receptor DBD
http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=pfam00105

Nuclear Receptor LBD
http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=pfam00104

It is particularly preferred if the steroid hormone receptor protein is estrogen receptor (ER).

In an alternative preferred embodiment it is preferred that the steroid hormone receptor protein is the androgen receptor (AR).

In a still more preferred embodiment, the binding of the nucleic acid binding portion to the target site is capable of being regulated by a ligand, which is preferably a molecule which does not naturally occur in the cell, more preferably does not occur naturally in an organism in which the cell is present. The ligand is preferably a small molecule, for example of less than 5000 daltons and is preferably bioavailable (for example when administered to a patient) and capable of entering the cell. Desirable properties for such a regulatory ligand will be well known to those skilled in the art.

It may be preferred that the ligand binding portion is a portion that does not naturally occur in a nucleic acid binding protein, for example a nuclear receptor DNA binding protein. For example, the ligand binding portion may be a modified ligand binding portion, for example as discussed in Doyle *et al* (2000) *Curr Opin Chem Biol* **4(1)**, 60-63 or Feil *et al* (1997) *Biochem Biophys Res Commun.* **237**, 752-757 (concerning a tamoxifen activated estrogen receptor ligand binding domain).

Thus, a preferred embodiment of the invention provides a method of suppressing the expression of a selected endogenous gene in a eukaryotic cell, the method comprising introducing into the cell (a) a molecule comprising a nucleic acid binding portion which binds to a site at or associated with the selected endogenous gene and a chromatin inactivation portion, or (b) a polynucleotide encoding said molecule, wherein the

ability of the molecule to bind to the site is capable of being regulated by a ligand, preferably a small molecule, and the method further comprises the step of exposing the cell to the ligand. It is preferred that the molecule does not comprise both the ligand binding domain and the DNA binding domain of the androgen receptor (AR). The molecule may therefore comprise a ligand binding portion.

It is believed that when the molecule comprises the estrogen receptor (ie the nucleic acid binding portion is provided as part of the estrogen receptor), then its ability to bind to the site is not capable of being regulated by an exogenous small molecule. Accordingly, in relation to this embodiment of the invention, the molecule does not comprise both the ligand binding domain and the DNA binding domain of the estrogen receptor (ER).

It is preferred in relation to this aspect of the invention that the molecule does not comprise both the ligand binding domain and nucleic acid binding domain of a nucleic acid binding protein occurring naturally in a cell, preferably occuring naturally in any cell. Thus, in a preferred embodiment, the nucleic acid binding portion and ligand binding portion are derivable from different polypeptides, for example from different nuclear receptor DNA binding protein, for example steroid hormone receptors.

In a particularly preferred embodiment, the ligand binding portion is derivable from the androgen receptor, for example as described in Example 1.

As also described in Example 1, the nucleic acid binding portion may be derivable from the estrogen receptor, but may alternatively be derived from another steroid hormone receptor, for example an "orphan" receptor (for example, a polypeptide identified as a steroid hormone receptor by sequence comparisons, but for which the nucleic acid binding specificity and/or ligand binding specificity are unknown). This embodiment may be particularly useful in investigating the biological role of "orphan" nuclear receptor DNA binding proteins, for example steroid hormone receptors.

It is preferred that the ligand is a non-naturally occurring small molecule. For example, the ligand may be the AR ligand R1881; thus, this ligand may be used when the molecule comprises the ligand binding domain from the AR as the ligand binding portion.

The ligand may be administered to the cell or a patient in order to regulate the suppression of the selected gene. This may be useful in a patient that it may be desirable to be able to regulate temporally expression of the selected gene, or at least the onset of repression or expression. It may also be useful when using the method to investigate the role of the gene (and or components of the molecule) in a cell or organism. For example, comparing cells/organisms in the absence of the ligand and/or molecule or polynucleotide encoding it, to cells/organisms in the presence of the ligand and/or molecule or polynucleotide encoding it, may be useful in distinguishing specific effects from non-specific effects, as well known to those skilled in the art.

Repression by histone deacetylation may also be reversed by using an inhibitor of histone deacetylase, for example Trichostatin A (TSA), Trapoxin or sodium butyrate (NaB), as known to those skilled in the art.

The ligand may promote or inhibit binding of the molecule to the nucleic acid; preferably it promotes binding to the nucleic acid. The ligand may inhibit binding by reducing the stability of the molecule. For example, ICI 164,384 is an anti-estrogen which binds to the ligand-binding domain of the estrogen receptor and decreases the half-life of the receptor, thereby reducing DNA binding (see Metzger *et al* (1995) Effect of antagonists on DNA binding properties of the human estrogen receptor *in vitro* and *in vivo*. *Mol Endocrinol* **9**, 579-591).

A further aspect of the invention provides a molecule, for example polypeptide, comprising a nucleic acid binding portion which binds to a site at or associated with a (preferably endogenous) eukaryotic gene, a ligand binding portion and a modulating portion which comprises a polypeptide or peptidomimetic which is capable of modulating covalent modification of nucleic acid or chromatin, for example a chromatin inactivation portion, wherein the ability of the molecule to bind to the site is capable of being regulated by binding of a ligand to the ligand binding portion, wherein the nucleic acid binding portion and ligand binding portion are not derivable from the same naturally occuring molecule, for example are not derivable from the same nuclear receptor DNA binding protein or steroid hormone receptor molecule. Preferences and definitions for the modifying or chromatin inactivation portion and nucleic acid binding portions are as discussed above in relation to the first aspect of the invention. For example, the modifying portion may comprise a methylase that methylates CpG dinucleotides, for example the methylase M.*Sss*I, a methyltransferase from *Spiroplasma* species (see Xu & Bestor (1997) *Nature Genet* **27**, 304-308).

It is particularly preferred that the ligand binding portion is or comprises the ligand binding domain of the androgen receptor, for example as described in Example 1.

Mangelsdorf *et al* (1995) *Cell* **83**, 835-839 reviews the steroid hormone receptor family, structure and nomenclature. Doyle *et al* (2000) *Curr Opin Chem Biol* **4(1)**, 60-63 discusses the ligand binding pocket of nuclear receptors and modification of ligand specificity using site-directed mutagenesis.

A still further aspect of the invention provides a method of suppressing the expression of a selected gene in a eukaryotic cell, the method comprising introducing into the cell (a) a molecule comprising a nucleic acid binding portion which binds to a site at or associated with the selected gene which site is present in a eukaryotic genome and a modifying portion which comprises a polypeptide or peptidomimetic which is capable of modulating covalent modification of nucleic acid or chromatin, for example a chromatin inactivation portion, or (b) a polynucleotide encoding said molecule, wherein the ability of the molecule to bind to the site is capable of being regulated by a ligand, preferably a small molecule, and the method further comprises the step of exposing the cell to the ligand, wherein the molecule does not comprise both the ligand binding domain and the DNA binding domain of the androgen receptor (AR). The molecule may therefore comprise a ligand binding portion.

In relation to this aspect of the invention the site present in a eukaryotic genome is a site which is at or associated with a selected gene or genes whose expression it is desirable to suppress or inactivate. It is preferred if

the site is a site which is naturally present in a eukaryotic genome. However, as is discussed in more detail below, the site may be one which has been engineered into the genome, or it may be a site associated with an inserted viral sequence. The site engineered into the genome to be in the vicinity of the gene whose expression is to be suppressed may be a site from the same species (but present elsewhere in the genome) or it may be a site present in a different species. By "genome" we include not only chromosomal DNA but other DNA present in the eukaryotic cell, such as DNA which has been introduced into the cell, for example plasmid or viral DNA. It is preferred if the nucleic acid binding portion can bind to chromosomal DNA or to RNA transcribed from chromosomal DNA.

For the avoidance of doubt, the site present in the eukaryote may be a naturally occurring site, or it may be a site which has been engineered to be there. The site need not be originally from the same or any other eukaryote. For example, it may be a bacterial repressor binding site which has been engineered to be present in the DNA of the eukaryotic cell, or it may be a mammalian nuclear receptor, for example steroid hormone receptor binding site which has been engineered into plant cells. However, it is preferred if the site to which the nucleic acid binding portion binds is naturally present in the eukaryotic cell and is present in its natural position in the genome.

Suitably the DNA binding portion may be all or a DNA-binding portion of an animal or plant DNA binding protein, or it may be all or a DNA binding portion of a bacterial or yeast DNA binding protein which has been engineered to bind to one or more sites in the plant or animal genome. Bacterial or yeast DNA binding proteins are less preferred and it is particularly preferred if the DNA binding protein does not contain a

DNA binding portion of wild-type *Saccharomyces cerevisiae* GAL4 or wild-type *Escherichia coli* LexA.

Preferences in relation to the previous aspect of the invention apply also to this aspect of the invention.

As noted, DNA binding proteins may be engineered so as to bind to a particular, selected target DNA sequence which is at or associated with a selected gene. In one embodiment of this aspect of the invention the DNA binding protein is one which has been engineered to bind to a site which is present in a mutant gene sequence within the plant or animal cell but is not present in the equivalent wild type sequence. For example, and as is discussed in more detail below, the engineered DNA binding portion may bind selectively to a dominant negative, mutated gene, such as a mutant oncogene and, upon binding, chromatin inactivation occurs and suppresses the expression of the mutant oncogene. Examples of oncogenes mutated in human cancer include RAS (*H-ras*) and *Bcl-10*.

Typically, the DNA binding portion and the modifying portion, for example chromatin inactivation portion, (and, where appropriate, the ligand binding portion) are fused such that the fusion polypeptide may be encoded by a nucleic acid molecule. Suitably, the DNA binding portion and the modifying or chromatin inactivation portion (and, where appropriate, the ligand binding portion) are joined so that each portion retains its respective activity such that the polypeptide may bind to a site present in a plant or animal genome (optionally in a ligand-regulatable, for example ligand-dependent, manner) and, upon binding, the modifying or chromatin inactivation portion is still able to inactivate nucleic acid or chromatin. The two (or three) portions may be joined directly, but they

may be joined by a linker peptide. Suitable linker peptides are those that typically adopt a random coil conformation, for example the polypeptide may contain alanine or proline or a mixture of alanine plus proline residues. Preferably, the linker contains between 10 and 100 amino acid residues, more preferably between 10 and 50 and still more preferably between 10 and 20. In any event, whether or not there is a linker between the portions of the polypeptide the polypeptide is able to bind its target DNA (optionally in a ligand regulatable, for example ligand-dependent, manner) and is able to inactivate nucleic acid or chromatin thereby selectively suppressing or inactivating gene expression.

A further aspect of the invention provides a polynucleotide encoding a polypeptide of the invention. In particular, the invention provides a polynucleotide wherein the nucleic acid binding portion, ligand binding portion and the modifying or chromatin inactivation portion are fused such that the fusion polypeptide is encoded by a single open reading frame of the polynucleotide. The polynucleotide may be DNA or RNA; DNA is preferred. DNA may or may not contain introns but, in any case, the polynucleotide encodes a polypeptide of the invention.

Polynucleotides which encode suitable nucleic acid binding portions, particularly DNA binding portions, or ligand binding portions are known in the art or can be readily designed from known sequences such as from known sequences contained in scientific publications or contained in nucleotide sequence databases such as the GenBank, EMBL and dbEST databases. Polynucleotides which encode suitable chromatin inactivation portions are known in the art or can readily be designed from known sequences and made. Polynucleotide sequences encoding various suitable modifying portions, for example chromatin inactivation portions are given

above in the references which refer to the polypeptides or are available from GenBank or EMBL or dbEST. A reference for PLZF is Chen *et al* (1993) *EMBO J.* **12**, 1161-1167.

Polynucleotides which encode suitable linker peptides can readily be designed from linker peptide sequences and made.

Thus, polynucleotides which encode the polypeptides of the invention (or polynucleotides as defined in relation to the first aspect of the invention) can readily be constructed using well known genetic engineering techniques.

A variety of methods have been developed to operably link polynucleotides, especially DNA, to other polynucleotides, including vectors, for example *via* complementary cohesive termini. For instance, complementary homopolymer tracts can be added to the DNA segment to be inserted to the vector DNA. The vector and DNA segment are then joined by hydrogen bonding between the complementary homopolymeric tails to form recombinant DNA molecules.

Synthetic linkers containing one or more restriction sites provide an alternative method of joining the DNA segment to vectors. The DNA segment, generated by endonuclease restriction digestion as described earlier, is treated with bacteriophage T4 DNA polymerase or *E. coli* DNA polymerase I, enzymes that remove protruding, 3'-single-stranded termini with their 3'-5'-exonucleolytic activities, and fill in recessed 3'-ends with their polymerising activities.

The combination of these activities therefore generates blunt-ended DNA

segments. The blunt-ended segments are then incubated with a large molar excess of linker molecules in the presence of an enzyme that is able to catalyse the ligation of blunt-ended DNA molecules, such as bacteriophage T4 DNA ligase. Thus, the products of the reaction are DNA segments carrying polymeric linker sequences at their ends. These DNA segments are then cleaved with the appropriate restriction enzyme and ligated to an expression vector that has been cleaved with an enzyme that produces termini compatible with those of the DNA segment.

Synthetic linkers containing a variety of restriction endonuclease sites are commercially available from a number of sources including International Biotechnologies Inc, New Haven, CN, USA.

A desirable way to modify the DNA encoding the polypeptide of the invention is to use the polymerase chain reaction as disclosed by Saiki *et al* (1988) *Science* **239**, 487-491. This method may be used for introducing the DNA into a suitable vector, for example by engineering in suitable restriction sites, or it may be used to modify the DNA in other useful ways as is known in the art.

In this method the DNA to be enzymatically amplified is flanked by two specific primers which themselves become incorporated into the amplified DNA. The said specific primers may contain restriction endonuclease recognition sites which can be used for cloning into expression vectors using methods known in the art.

Methods of joining a polynucleotide to a nucleic acid vector are, of course, applicable to joining any polynucleotides.

The DNA (or in the case of retroviral vectors, RNA) is then expressed in a suitable host to produce a polypeptide of the invention. Thus, the DNA encoding the polypeptide of the invention may be used in accordance with known techniques, appropriately modified in view of the teachings contained herein, to construct an expression vector, which is then used to transform an appropriate host cell for the expression and production of the polypeptide of the invention. Such techniques include those disclosed in US Patent Nos. 4,440,859 issued 3 April 1984 to Rutter *et al*, 4,530,901 issued 23 July 1985 to Weissman, 4,582,800 issued 15 April 1986 to Crowl, 4,677,063 issued 30 June 1987 to Mark *et al*, 4,678,751 issued 7 July 1987 to Goeddel, 4,704,362 issued 3 November 1987 to Itakura *et al*, 4,710,463 issued 1 December 1987 to Murray, 4,757,006 issued 12 July 1988 to Toole, Jr. *et al*, 4,766,075 issued 23 August 1988 to Goeddel *et al* and 4,810,648 issued 7 March 1989 to Stalker, all of which are incorporated herein by reference.

The DNA (or in the case of retroviral vectors, RNA) encoding the polypeptide of the invention may be joined to a wide variety of other DNA sequences for introduction into an appropriate host. The companion DNA will depend upon the nature of the host, the manner of the introduction of the DNA into the host, and whether episomal maintenance or integration is desired.

Generally, the DNA is inserted into an expression vector, such as a plasmid, in proper orientation and correct reading frame for expression. If necessary, the DNA may be linked to the appropriate transcriptional and translational regulatory control nucleotide sequences recognised by the desired host, although such controls are generally available in the expression vector. The vector is then introduced into the host through

standard techniques. Generally, not all of the hosts will be transformed by the vector. Therefore, it will be necessary to select for transformed host cells. One selection technique involves incorporating into the expression vector a DNA sequence, with any necessary control elements, that codes for a selectable trait in the transformed cell, such as antibiotic resistance. Alternatively, the gene for such selectable trait can be on another vector, which is used to co-transform the desired host cell.

Host cells that have been transformed by the recombinant DNA of the invention are then cultured for a sufficient time and under appropriate conditions known to those skilled in the art in view of the teachings disclosed herein to permit the expression of the polypeptide, which can then be recovered.

Many expression systems are known, including bacteria (for example *E. coli* and *Bacillus subtilis*), yeasts (for example *Saccharomyces cerevisiae*), filamentous fungi (for example *Aspergillus*), plant cells, animal cells and insect cells.

The vectors include a prokaryotic replicon, such as the ColE1 *ori*, for propagation in a prokaryote, even if the vector is to be used for expression in other, non-prokaryotic, cell types. The vectors can also include an appropriate promoter such as a prokaryotic promoter capable of directing the expression (transcription and translation) of the genes in a bacterial host cell, such as *E. coli*, transformed therewith.

A promoter is an expression control element formed by a DNA sequence that permits binding of RNA polymerase and transcription to occur. Promoter sequences compatible with exemplary bacterial hosts are

typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention. It is preferred that the promoter is one which can be regulated. It is particularly preferred if the promoter is an inducible promoter which can be selectively induced at an appropriate time once the vector has been introduced into the eukaryotic cell. It will be appreciated that upon induction, the polypeptide of the invention may be expressed in the cell and exert its effect. In this situation, induction of expression of the polypeptide of the invention leads to suppression of the targeted gene. Inducible promoters are known in the art for many eukaryotic cells including plant and animal cells. These include heat-shock-, glucocorticoid-, oestradiol-, and metal-inducible promoter systems.

Typical prokaryotic vector plasmids are pUC18, pUC19, pBR322 and pBR329 available from Biorad Laboratories, (Richmond, CA, USA) and p*Trc*99A and pKK223-3 available from Pharmacia, Piscataway, NJ, USA.

A typical mammalian cell vector plasmid is pSVL available from Pharmacia, Piscataway, NJ, USA. This vector uses the SV40 late promoter to drive expression of cloned genes, the highest level of expression being found in T antigen-producing cells, such as COS-1 cells.

An example of an inducible mammalian expression vector is pMSG, also available from Pharmacia. This vector uses the glucocorticoid-inducible promoter of the mouse mammary tumour virus long terminal repeat to drive expression of the cloned gene.

Useful yeast plasmid vectors are pRS403-406 and pRS413-416 and are generally available from Stratagene Cloning Systems, La Jolla, CA 92037,

USA. Plasmids pRS403, pRS404, pRS405 and pRS406 are Yeast Integrating plasmids (YIps) and incorporate the yeast selectable markers *HIS3, TRP1, LEU2* and *URA3*. Plasmids pRS413-416 are Yeast Centromere plasmids (YCps).

Plant transformation vectors are well known in the art. For example, vectors for *Agrobacterium*-mediated transformation are available from the Centre for the Application of Molecular Biology to International Agriculture, GPO Box 3200, Canberra, ACT 2601, Australia (cambia@cambia.org.au).

The present invention also relates to a host cell transformed with a polynucleotide vector construct of the present invention. The host cell can be either prokaryotic or eukaryotic. Bacterial cells are preferred prokaryotic host cells and typically are a strain of *E. coli* such as, for example, the *E. coli* strains DH5 available from Bethesda Research Laboratories Inc., Bethesda, MD, USA, and RR1 available from the American Type Culture Collection (ATCC) of Rockville, MD, USA (No ATCC 31343). Preferred eukaryotic host cells include plant, yeast, insect and mammalian cells, preferably vertebrate cells such as those from a mouse, rat, monkey or human fibroblastic and kidney cell lines. Yeast host cells include YPH499, YPH500 and YPH501 which are generally available from Stratagene Cloning Systems, La Jolla, CA 92037, USA. Preferred mammalian host cells include Chinese hamster ovary (CHO) cells available from the ATCC as CCL61, NIH Swiss mouse embryo cells NIH/3T3 available from the ATCC as CRL 1658, monkey kidney-derived COS-1 cells available from the ATCC as CRL 1650 and 293 cells which are human embryonic kidney cells. Preferred insect cells are Sf9 cells which can be transfected with baculovirus expression vectors.

Protoplasts for transformation are typically generated as required by methods known in the art. Plant cell lines are not generally available. However, one cell line which is commonly used is the Bright Yellow 2 cell line from tobacco (BY2; Mu *et al* (1997) *Plant Mol. Biol.* **34**, 357-362).

Transformation of appropriate cell hosts with a DNA construct of the present invention is accomplished by well known methods that typically depend on the type of vector used. With regard to transformation of prokaryotic host cells, see, for example, Cohen *et al* (1972) *Proc. Natl. Acad. Sci. USA* **69**, 2110 and Sambrook *et al* (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. Transformation of yeast cells is described in Sherman *et al* (1986) *Methods In Yeast Genetics, A Laboratory Manual*, Cold Spring Harbor, NY. The method of Beggs (1978) *Nature* **275**, 104-109 is also useful. With regard to vertebrate cells, reagents useful in transfecting such cells, for example calcium phosphate and DEAE-dextran or liposome formulations, are available from Stratagene Cloning Systems, or Life Technologies Inc., Gaithersburg, MD 20877, USA. With regard to plant cells and whole plants three plant transformation approaches are typically used (J. Draper and R. Scott in D. Grierson (ed.), "Plant Genetic Engineering", Blackie, Glasgow and London, 1991, vol. 1, pp 38-81):

i) *Agrobacterium*-mediated transformation, using the Ti plasmid of *A. tumefaciens* and the Ri plasmid of *A. rhizogenes* (P. Armitage, R. Walden and J. Draper in J. Draper, R. Scott, P. Armitage and R. Walden (eds.), "Plant Genetic Transformation and Expression - A Laboratory Manual", Blackwell Scientific Publications, Oxford, 1988, pp 1-67; R.J. Draper, R.

Scott and J. Hamill *ibid.*, pp 69-160);

Agrobacterium-mediated transformation is also described in Hooykaas & Schilperoot (1992) *Plant Mol. Biol.* **19**, 15-38; Zupan & Zambryski (1995) *Plant Physiol.* **107**, 1041-1047; and Baron & Zambryski (1996) *Curr. Biol.* **6**, 1567-1569.

ii) DNA-mediated gene transfer, by polyethylene glycol-stimulated DNA uptake into protoplasts, by electroporation, or by microinjection of protoplasts or plant cells (J. Draper, R. Scott, A. Kumar and G. Dury, *ibid.*, pp 161-198). Direct gene transfer into protoplasts is also described in Neuhaus & Spangenberg (1990) *Physiol. Plant* **79**, 213-217; Gad *et al* (1990) *Physiol. Plant* **79**, 177-183; and Mathur & Koncz (1998) *Method Mol. Biol.* **82**, 267-276;

iii) transformation using particle bombardment (D. McCabe and P. Christou, *Plant Cell Tiss. Org. Cult.*, **3**, 227-236 (1993); P. Christou, *Plant J.*, **3**, 275-281 (1992)).

Some species are amenable to direct transformation, avoiding a requirement for tissue or cell culture (Bechtold *et al* (1993) *Life Sciences*, C.R. Acad. Sci. Paris **316**, 1194-1199).

Agrobacterium-mediated transformation is generally less effective for monocotyledonous plants for which approaches ii) and iii) are therefore preferred. However, *Agrobacterium* is capable of transferring DNA to some monocotyledenous plants if tissues containing "competent" cells are infected (see Hiei *et al* (1997) *Plant Mol. Biol.* **35**, 205-218). In all approaches a suitable selection marker, such as kanamycin- or herbicide-

resistance, is preferred or alternatively a screenable marker ("reporter") gene, such as ∃-glucuronidase or luciferase (see J. Draper and R. Scott in D. Grierson (ed.), "Plant Genetic Engineering", Blackie, Glasgow and London, 1991, vol. 1 pp 38-81).

Electroporation is also useful for transforming and/or transfecting cells and is well known in the art for transforming yeast cell, bacterial cells, insect cells, vertebrate cells and some plant cells (eg barley cells, see Lazzeri (1995) *Methods Mol. Biol.* **49**, 95-106).

For example, many bacterial species may be transformed by the methods described in Luchansky *et al* (1988) *Mol. Microbiol.* **2**, 637-646 incorporated herein by reference. The greatest number of transformants is consistently recovered following electroporation of the DNA-cell mixture suspended in 2.5X PEB using 6250V per cm at 25:FD.

Methods for transformation of yeast by electroporation are disclosed in Becker & Guarente (1990) *Methods Enzymol.* **194**, 182.

Successfully transformed cells, ie cells that contain a DNA construct of the present invention, can be identified by well known techniques. For example, cells resulting from the introduction of an expression construct of the present invention can be grown to produce the polypeptide of the invention. Cells can be harvested and lysed and their DNA content examined for the presence of the DNA using a method such as that described by Southern (1975) *J. Mol. Biol.* **98**, 503 or Berent *et al* (1985) *Biotech.* **3**, 208. Alternatively, the presence of the protein in the supernatant can be detected using antibodies as described below.

In addition to directly assaying for the presence of recombinant DNA, successful transformation can be confirmed by well known immunological methods when the recombinant DNA is capable of directing the expression of the polypeptide. For example, cells successfully transformed with an expression vector produce polypeptides displaying appropriate antigenicity. Samples of cells suspected of being transformed are harvested and assayed for the protein using suitable antibodies.

Thus, in addition to the transformed host cells themselves, the present invention also contemplates a culture of those cells, preferably a monoclonal (clonally homogeneous) culture, or a culture derived from a monoclonal culture, in a nutrient medium.

In relation to plants, it is envisaged that the invention includes single cell derived cell suspension cultures, isolated protoplasts or stable transformed plants. In the latter case it is preferred if the polypeptide of the invention (or molecule as defined in relation to the first aspect of the invention) is expressed using an inducible promoter system to avoid potentially lethal effects of gene down-regulation during regeneration of homozygous plants. Alternatively or in addition, the binding of the polypeptide to the site may be regulatable by a ligand, as discussed above, and regeneration of homozygous plants performed under conditions which do not promote binding of the polypeptide to the site.

Although the molecules, polypeptides or polynucleotides of the invention (or as defined in relation to the invention) may be introduced into any suitable host cell, it will be appreciated that they are primarily designed to be effective in appropriate animal or plant cells, particularly those that have one or more sites within their DNA to which the molecule (for

example polypeptide of the invention) may bind.

Thus, the animal of plant cells which contain a molecule, polypeptide (or polynucleotide) of the invention (or which have been subjected to a method of the invention) whose presence suppresses the expression of a particular gene, or the animals or plants containing these cells, may be considered to have the gene "knocked out" in the sense that it can no longer be expressed. The DNA methylation or chromatin inactivation by histone deacetylation may be irreversible.

It will be readily appreciated that introduction of a molecule or polypeptide of the invention, or as defined in relation to the invention, into an animal or plant cell, or introduction and expression of a polynucleotide encoding such a polypeptide in an animal or plant cell, will allow targeting of the molecule or polypeptide to an appropriate binding site within the DNA (and which is bound by the DNA-binding portion of the molecule orpolypeptide) and allow for the DNA or chromatin at or associated with the target binding site to be inactivated so as to lead to suppression or inactivation of gene expression. Typically, the molecule or polypeptide of the invention is selected so that it targets a selected gene. Thus, suitably, the targeted gene has a site which is bound by the DNA binding portion of the molecule or polypeptide associated with it. The site which is so bound may be within the gene itself, for example within an intron or within an exon of the gene; or it may be in a region 5N of the transcribed portion of the gene, for example within or adjacent to a promoter or enhancer region; or it may be in a region 3N of the transcribed portion of the gene.

Genes regulated by oestrogen receptor (ER) include the progesterone receptor (PR) gene and the PS2 (trefoil related protein) gene. Thus, the

method of the invention may be used to inactivate the PR gene or the PS2 gene when the DNA binding portion of the compound of the invention is at least the DNA-binding portion of ER. Anti-oestrogen therapy is used in the treatment of breast cancer. The full repertoire of oestrogen regulated genes involved in breast cancer is presently unknown. It is generally considered that anti-oestrogen therapy results in the altered expression of key oestrogen regulated genes involved in breast cancer cell growth and transformation. The methods of the invention described below may provide an alternative, potentially more effective, way of regulating the expression (particularly inhibiting) of oestrogen-responsive genes. It may be that for certain DNA binding portions, in a given plant or animal cell there is only one target site and the expression of only one gene is suppressed by the modifying portion or chromatin inactivation portion. However, there may be more than one target site and introduction of a molecule or polypeptide (or polynucleotide) of the invention may lead to suppression of expression of a number of genes.

The ability to suppress the expression of a selected gene is useful in many areas of biology.

Typically, when the gene whose expression is suppressed is in an animal cell, the animal cell is a cell within an animal and the method of the invention is used to suppress the expression of a selected gene in an animal. For the avoidance of doubt, animal in this context includes human; however, the animal may be non-human. Examples of particular uses in animal cells include allele-specific inactivation of oncogenic proteins such as mutant *Ras* and mutant *Bcl-10*; inhibition of oestrogen receptor regulated gene expression in breast cancer; inhibition of androgen receptor; inhibition of genes of interest for developmental studies;

inhibition of genes for developing transgenic models of human diseases; inhibition of genes involved in tissue modelling, as found in cancer and wound healing; and identification of function(s) for genes and/or nucleic acid binding proteins/ligand receptors.

Also typically, the plant cell is a cell within a plant and the method of the invention is used to suppress the expression of a selected gene in a plant.

In one embodiment, the method of the invention is used to suppress the expression of socially or environmentally unacceptable or undesirable genes in commercially engineered transgenic plants. Such genes may include, for example, antibiotic or herbicide selectable marker genes. In this embodiment, the gene in the transgenic plant is targeted for silencing.

In a further embodiment of the invention novel plant architecture or floral morphology may be achieved by targeting some known homeotic genes involved in these developmental pathways.

Suitably, the method of the invention is used to suppress or inactivate the expression of a gene whose expression it is desirable to suppress or inactivate. Such genes include oncogenes, viral genes including genes present in proviral genomes and so the method in relation to animals may constitute a method of medical treatment. Oncogenes may be overexpressed in certain cancers and it may be desirable to suppress their expression. Some oncogenes are oncogenic by virtue of having an activating mutation. Using the method of the invention the selective suppression of expression of a mutant oncogene may be achieved using a DNA binding portion that selectively binds to the mutant oncogene sequence and wherein the modifying or chromatin inactivation portion

inactivates the DNA or chromatin in which the oncogene resides or with which it is associated so that expression of the mutant oncogene is suppressed. Suppression of oncogene overexpression or of mutant (especially activated) oncogene expression is generally desirable in treating cancers in which the oncogenes play a role. Mutant oncogenes which may be targeted by the method of the invention include *Ras* and *Bcl-10*. These may be targeted by engineered DNA binding proteins capable of recognising the mutated genes in a sequence specific manner.

The expression of viral genes in an animal or plant cell is generally undesirable since this expression is often associated with pathogenesis. The nucleic acid of certain viruses may be formed into chromatin and the expression of such viral genes may be controlled by modification of this chromatin. For example, retroviral proviruses (ie DNA copies of retroviral RNAs) are often incorporated into animal and plant genomes where they become part of the chromatin, for example, integrated HIV provirus and integrated human papillomavirus. *Gypsy* and *Copia*-like retrotransposons appear to be widely distributed in the plant kingdom. *Copia*-like retrotransposons, or at least their reverse transcriptase domains, appear broadly distributed in higher plants while the *Gypsy*-like elements (which share their organisation with the retroviruses but lack retroviral envelope domains) are less abundant (Suoniemi *et al* (1998) *Plant J.* **13**, 699-705). Integration of viral DNA into the plant genome has been demonstrated for geminiviral DNA into the tobacco nuclear genome (Bejarano *et al* (1996) *Proc. Natl. Acad. Sci. USA* **93**, 759-764). Potential retroviruses have also recently been described in plants (Wright & Voytus (1998) *Genetics* **149**, 703-715). Using a method of the invention the selective suppression of expression of a viral gene may be achieved. Engineered DNA binding proteins, or RNA binding proteins,

such as HIV *tat* protein, may be used to target a modifying or chromatin inactivation portion and lead to proviral genome inactivation by binding to nascent genomic RNA transcripts, achieving, for example, DNA methylation or histone deacetylation by proximity.

Certain genetic diseases are caused by dominant mutations, such as achondroplasia. Suppression of expression of the mutant allele may be useful in treating these diseases. Using a method of the invention the selective suppression of expression of the mutant allele may be achieved using a DNA binding portion that selectively binds to the mutant allele sequence and wherein the modifying or chromatin inactivation portion inactivates the DNA or chromatin in which the mutant allele resides or with which it is associated so that expression of the mutant allele is suppressed.

These methods of the invention typically and preferably involve the transfer of a polynucleotide encoding said polypeptide into an animal or plant cell.

Gene transfer systems known in the art may be useful in the practice of the methods of the present invention, for example in which the polynucleotide of the invention is introduced into a cell either within or outwith an animal body. Such an introduction of a polynucleotide may be therapeutically useful and constitutes a form of gene therapy. These include viral and nonviral transfer methods. A number of viruses have been used as gene transfer vectors, including papovaviruses, eg SV40 (Madzak *et al* (1992) *J. Gen. Virol.* **73**, 1533-1536), adenovirus (Berkner (1992) *Curr. Top. Microbiol. Immunol.* **158**, 39-61; Berkner *et al* (1988) *BioTechniques* **6**, 616-629; Gorziglia and Kapikian (1992) *J. Virol.* **66**,

4407-4412; Quantin *et al* (1992) *Proc. Natl. Acad. Sci. USA* **89**, 2581-2584; Rosenfeld *et al* (1992) *Cell* **68**, 143-155; Wilkinson *et al* (1992) *Nucleic Acids Res.* **20**, 2233-2239; Stratford-Perricaudet *et al* (1990) *Hum. Gene Ther.* **1**, 241-256), vaccinia virus (Moss (1992) *Curr. Top. Microbiol. Immunol.* **158**, 25-38), adeno-associated virus (Muzyczka (1992) *Curr. Top. Microbiol. Immunol.* **158**, 97-123; Ohi *et al* (1990) *Gene* **89**, 279-282), herpes viruses including HSV and EBV (Margolskee (1992) *Curr. Top. Microbiol. Immunol.* **158**, 67-90; Johnson *et al* (1992) *J. Virol.* **66**, 2952-2965; Fink *et al* (1992) *Hum. Gene Ther.* **3**, 11-19; Breakfield and Geller (1987) *Mol. Neurobiol.* **1**, 337-371; Freese *et al* (1990) *Biochem. Pharmacol.* **40**, 2189-2199), and retroviruses of avian (Brandyopadhyay and Temin (1984) *Mol. Cell. Biol.* **4**, 749-754; Petropoulos *et al* (1992) *J. Virol.* **66**, 3391-3397), murine (Miller (1992) *Curr. Top. Microbiol. Immunol.* **158**, 1-24; Miller *et al* (1985) *Mol. Cell. Biol.* **5**, 431-437; Sorge *et al* (1984) *Mol. Cell. Biol.* **4**, 1730-1737; Mann and Baltimore (1985) *J. Virol.* **54**, 401-407; Miller *et al* (1988) *J. Virol.* **62**, 4337-4345), and human origin (Shimada *et al* (1991) *J. Clin. Invest.* **88**, 1043-1047; Helseth *et al* (1990) *J. Virol.* **64**, 2416-2420; Page *et al* (1990) *J. Virol.* **64**, 5370-5276; Buchschacher and Panganiban (1992) *J. Virol.* **66**, 2731-2739). To date most human gene therapy protocols have been based on disabled murine retroviruses.

Nonviral gene transfer methods known in the art include chemical techniques such as calcium phosphate coprecipitation (Graham and van der Eb (1973) *Virology* **52**, 456-467; Pellicer *et al* (1980) *Science* **209**, 1414-1422); mechanical techniques, for example microinjection (Anderson *et al* (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5399-5403; Gordon *et al,* 1980; Brinster *et al* (1981) *Cell* **27**, 223-231; Constantini and Lacy (1981) *Nature* **294**, 92-94); membrane fusion-mediated transfer *via* liposomes

(Felgner *et al* (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7413-7417; Wang and Huang (1989) *Biochemistry* **28**, 9508-9514; Kaneda *et al* (1989) *J. Biol. Chem.* **264**, 12126-12129; Stewart *et al* (1992) *Hum. Gene Ther.* **3**, 267-275; Nabel *et al,* 1990; Lim *et al* (1992) *Circulation* **83**, 2007-2011); and direct DNA uptake and receptor-mediated DNA transfer (Wolff *et al* (1990) *Science* **247**, 1465-1468; Wu *et al* (1991) *J. Biol. Chem.* **266**, 14338-14342; Zenke *et al* (1990) *Proc. Natl. Acad. Sci. USA* **87**, 3655-3659; Wu *et al,* 1989b; Wolff *et al* (1991) *BioTechniques* **11**, 474-485; Wagner *et al,* 1990; Wagner *et al* (1991) *Proc. Natl. Acad. Sci. USA* **88**, 4255-4259; Cotten *et al* (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4033-4037; Curiel *et al* (1991a) *Proc. Natl. Acad. Sci. USA* **88**, 8850-8854; Curiel *et al* (1991b) *Hum. Gene Ther.* **3**, 147-154). Viral-mediated gene transfer can be combined with direct *in vivo* gene transfer using liposome delivery, allowing one to direct the viral vectors to the tumour cells and not into the surrounding nondividing cells. Alternatively, the retroviral vector producer cell line can be injected into tumours (Culver *et al* (1992) *Science* **256**, 1550-1552). Injection of producer cells would then provide a continuous source of vector particles. This technique has been approved for use in humans with inoperable brain tumours.

Other suitable systems include the retroviral-adenoviral hybrid system described by Feng *et al* (1997) *Nature Biotechnology* **15**, 866-870, or viral systems with targeting ligands such as suitable single chain Fv fragments.

In an approach which combines biological and physical gene transfer methods, plasmid DNA of any size is combined with a polylysine-conjugated antibody specific to the adenovirus hexon protein, and the resulting complex is bound to an adenovirus vector. The trimolecular complex is then used to infect cells. The adenovirus vector permits

efficient binding, internalization, and degradation of the endosome before the coupled DNA is damaged.

Liposome/DNA complexes have been shown to be capable of mediating direct *in vivo* gene transfer. While in standard liposome preparations the gene transfer process is nonspecific, localized *in vivo* uptake and expression have been reported in tumour deposits, for example, following direct *in situ* administration (Nabel (1992) *Hum. Gene Ther.* 3, 399-410).

Gene transfer techniques which target DNA directly to a target cell or tissue, is preferred. Receptor-mediated gene transfer, for example, is accomplished by the conjugation of DNA (usually in the form of covalently closed supercoiled plasmid) to a protein ligand *via* polylysine. Ligands are chosen on the basis of the presence of the corresponding ligand receptors on the cell surface of the target cell/tissue type. These ligand-DNA conjugates can be injected directly into the blood if desired and are directed to the target tissue where receptor binding and internalization of the DNA-protein complex occurs. To overcome the problem of intracellular destruction of DNA, coinfection with adenovirus can be included to disrupt endosome function.

It may be advantageous if the polypeptide of the invention, or as defined in relation to the first aspect of the invention is expressed in the target cell using an inducible promoter. Examples of suitable inducible promoters include those that can be induced by heat shock, glucocorticoids, oestradiol and metal ions.

Preferably, the method of suppressing the expression of a selected gene is used to suppress expression of a gene in a human cell; in one particularly preferred embodiment the human cell is within a human body.

However, the method of the invention may involve the modification of animal cells (including human cells) outside of the body of an animal (ie an *ex vivo* treatment of the cells) and the so modified cells may be reintroduced into the animal body.

The method may also be used on cells outside an animal body, ie *in vitro*, for example on a stable cell line.

From the foregoing, it will be appreciated that the method of the invention may be useful to suppress the activity of a plurality of selected genes. In particular, the method of the invention may be used to suppress the activity of a group of genes whose expression is controlled, at least to a large extent, by a single transcription factor. For example, the method may be used to suppress oestrogen-regulated genes as is described in more detail in the Examples.

A further aspect of the invention provides use of a molecule, for example polypeptide, of the invention (ie comprising a nucleic acid binding portion which binds to a site at or associated with an endogenous eukaryotic gene, a ligand binding portion and a modifying or chromatin inactivation portion, wherein the ability of the molecule to bind to the site is capable of being regulated by binding of a ligand to the ligand binding portion, wherein the nucleic acid binding portion and ligand binding portion are not derivable from the same naturally occuring molecule, for example are not derivable from the same steroid hormone receptor molecule), or

polynucleotide of the invention, in the manufacture of an agent for suppressing the expression of the selected gene in a eukaryotic cell.

It will be appreciated that it is particularly preferred if the molecule, polypeptide or polynucleotide is used in the preparation of a medicament for suppressing the expression of a selected gene in an animal. For the avoidance of doubt, by "animal" we include human.

A further aspect of the invention provides a method of treating a patient in need of suppression of the expression of a selected gene, the method comprising administering to the patient an effective amount of a molecule, for example polypeptide, of the invention, or polynucleotide of the invention.

Preferably the patient is also administered a ligand molecule which is capable of modulating the binding of the molecule or polypeptide (or encoded polypeptide if a polynucleotide is administered) to the site. Such administration may take place before, simultaneously with, or after administration of the molecule, polypeptide or polynucleotide. It may take place by the same route, or by a different route. For example, the ligand molecule may be administered orally. For example, possible ligands CPA, casodex, testosterone and Dihydrotestosterone (and similar compounds) may be administered orally.

It will be appreciated that suppression of the expression of a selected gene is useful where the expression or overexpression of the selected gene is undesirable and contributes to a disease state in the patient. Examples of undesirable expression of a gene include the expression of certain activated oncogenes in cancer.

Suppression of the expression of the ER upregulated genes is desirable in the treatment of breast cancer. Similarly, suppression of the expression of the androgen receptor (AR)-regulated genes is desirable in the treatment of prostate cancer.

Further aspects of the invention provides use of a molecule, polypeptide or polynucleotide of the invention in the manufacture of a medicament for suppressing the expression of a selected gene in a patient in need of such suppression.

Still further aspects of the invention provides a molecule, polypeptide or polynucleotide of the invention for use in medicine. Thus, the molecule, polypeptide or polynucleotide are packaged and presented for use in medicine.

Yet still further aspects of the invention provide a pharmaceutical composition comprising a molecule, polypeptide or polynucleotide of the invention and a pharmaceutically acceptable carrier.

By "pharmaceutically acceptable" is included that the formulation is sterile and pyrogen free. Suitable pharmaceutical carriers are well known in the art of pharmacy.

A further aspect of the invention provides a method for designing a molecule for suppressing expression of a selected endogenous gene in a eukaryotic cell, the method comprising

(1) identifying a site at or associated with the selected endogenous gene

(2) identifying or designing a nucleic acid binding portion which binds to, or is predicted to bind to, the site (or a polynucleotide having or comprising the nucleotide sequence of the site)

(3) preparing a molecule comprising the nucleic acid binding portion and a modifying (as defined above) or chromatin inactivation portion, or a polynucleotide encoding said molecule.

The method preferably further comprises the steps of

(4) testing the affinity and/or specificity of binding of the nucleic acid binding portion to the site and/or a polynucleotide having or comprising the nucleotide sequence of the site; and/or

(5) testing the affinity and/or specificity of binding of the molecule to the site and/or a polynucleotide having or comprising the nucleotide sequence of the site; and/or

(6) testing the efficacy of the molecule or polynucleotide in suppressing the expression of the gene and/or of a reporter gene comprising the nucleotide sequence of the site.

The nucleic acid binding portion may be identified or designed using methods known to those skilled in the art, for example using algorithms or for design of zinc finger proteins for binding to particular nucleotide sequences (see, for example, WO96/06166, WO00/42219, WO98/53058, WO98/53059, WO98/53060) or using library screening, for example using phage display techniques (see, for example WO98/53057, WO00/27878). Protein array chip techniques and nucleic acid chip techniques may be used. Such techniques are reviewed in, for example, the following references:

DNA array:

David J Duggan, Michael Bittner, Yidong Chen, Paul Meltzer & Jeffrey M. Trent (1999) Expression profiling using cDNA microarrays. *Nat. Genet.* 21 (suppl) 10 - 14

Patrick O. Brown & David Botstein (1999) Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* 21 (suppl) 33 - 37

Christine Debouck & Peter N. Goodfellow (1999) DNA microarrays in drug discovery and development. *Nat. Genet.* 21 (suppl) 48 - 50.

proteomics:

Mann M, Hendrickson RC, Pandey A. (2001) Analysis of proteins and proteomes by mass spectrometry. *Annu Rev Biochem.* 70:437-473.

Pandey A, Mann M. (2000) Proteomics to study genes and genomes. *Nature*. 405:837-846.

Methods for preparing the molecule, for example polypeptide, once a nucleic acid binding domain has been identified or designed, are described above and will be apparent to those skilled in the art. Suitable methods for testing the affinity and/or specificity of binding will also be known to those skilled in the art, and exemplary techniques are discussed in the Examples. Methods for preparing and using reporter gene constructs will also be known to those skilled in the art and are discussed in the Examples.

A further aspect of the invention provides a method for preparing a molecule for use in identifying a eukaryotic gene or genes regulated in a

cell by a putative nucleic acid binding protein, the method comprising the steps of

(1) preparing (a) a molecule comprising (i) at least the predicted nucleic acid binding portion of the putative nucleic acid binding protein and (ii) a modifying or chromatin inactivation portion and, optionally (iii) a ligand binding domain, or (b) a polynucleotide encoding said molecule.

The molecule may be useful in a method comprising the steps of

(1) introducing into the cell (a) the said molecule or (b) the said polynucleotide, and optionally

(2) determining the effect of step (1) on the cell.

The invention further provides a method for identifying a eukaryotic gene or genes regulated in a cell by a putative nucleic acid binding protein, the method comprising the steps of

(1) preparing (a) a molecule comprising (i) at least the predicted nucleic acid binding portion of the putative nucleic acid binding protein and (ii) a modifying or chromatin inactivation portion and, optionally (iii) a ligand binding domain, or (b) a polynucleotide encoding said molecule.

(2) introducing into the cell (a) the said molecule or (b) the said polynucleotide, and optionally

(3) determining the effect of step (2) on the cell.

For example, the effect of the molecule or polynucleotide on the cell may be determined by comparing a cell prior to the introduction with the cell after the introduction (for example using single cell imaging or other analysis); or by comparing a cell (or population of cells) to which the molecule or polypeptide has been introduced with a cell (or population of cells) to which the molecule or polypeptide has not been introduced; or, if

the molecule comprises a ligand binding domain, by comparing a cell (or population of cells) to which the ligand and the molecule or polypeptide have been introduced with a cell (or population of cells) to which the molecule or polypeptide has been introduced but to which the ligand has not been introduced.

Comparisons may be made using methods well known to those skilled in the art, for example using RNA expression analysis, gene profiling, proteomic analysis (for example using high throughput techniques, references for which are given above) and/or phenotypic analysis.

All publications and patent documents referred to herein are hereby incorporated by reference.

The invention will now be described in more detail with reference to the following Figures and Examples:

Figure 1: PLZF-AR inhibits an androgen-responsive reporter gene in a ligand-dependent manner.

(A) Schematic representation of RARα, AR and ERα, showing the positions of the DNA binding domain (DBD) and the ligand binding domain (LBD). Also shown are transcription activation functions AF-1 and AF-2.

(B-E) COS-1 cells were transiently transfected with a CAT reporter gene containing an androgen response element and the full length human AR, AR lacking the LBD (ARΔLBD), PLZF-AR or PANER as shown. As all of the expression plasmids were cloned into the mammalian expression plasmid pSG5, pSG5 was transfected as a control (-). R1881, a synthetic androgen was added to a final concentration of 1 nM, 17β-estradiol

(10nM). All other ligands were added at 100 nM. As all ligands were prepared in ethanol, an equal volume of ethanol was added to the no ligand controls. Results represent at least three independent experiments. In each experiment AR activity in the presence of R1881 was taken as 100%. All other activities are shown relative to this.

Figure 2: Immunoblotting for PLZF-AR. Cell lysates made from COS-1 cells transiently transfected with an expression plasmid encoding the androgen receptor (AR, lane 2), ARΔLBD (lane 3), PLZF-AR (lane 40, PLZF-AR-ER (PANER, lane 5), PLZF (lane 6) or the parental expression plasmid pSG5 (lane 1) were separated by 10% SDS-PAGE and immunoblotted using antibodies to the N-terminal or C-terminal portions of AR or using antibodies against PLZF.

Figure 3: The LnCAP prostate cell line was transiently transfected with the adenoviral shuttle vector pADTRack-CMV encoding PLZF-AR 9pAD-PLZF-AR). Transfected cells were FACS sorted for GFP expression, which is encoded by the pADTrack-CMV vector. GFP-positive, FACS sorted cells were cultured in the presence or absence of the AR agonist R1881. After 72 hours, culture media were collected and the amount of the androgen-regulated protein PSA determined by immunoassay. PSA concentrations in ng/ml are shown. Control cell transfected with the parent vector (pADTrack-CMV) were treated in the same way. (-) refers to determination of the PSA level of the medium used to culture the cells (lane 5). Addition of R1881 resulted in a 25-fold increase in levels of secreted PSA in cells transfected with pADTrack-CMV vector alone (lanes 3,4). By contrast PSA levels in pAD-PLZF-AR transfected cells were raised 5-fold, demonstrating a 5-fold decrease in PSA levels, relative to the pADTRack-CMV control (compare lanes 2 and 4). At the time of

assaying for PSA, approximately 70% of the FACS sorted cells remained GFP-positive, in both pAD_PLZF-AR and pADTrack-CMV transfected cultures. This suggests that induced PSA activity in pAD-PLZF-AR transfected cultures arises from cells that have lost construct DNA.

Figure 4: PLZF-AR represses endogenous Androgen-regulated gene expression. The Androgen Receptor (AR)-positive human T47D cell line was infected with an adenovirus encoding PLZF-AR and green fluorescnet protein (GFP) or the parental adenovirus (GFP). Infected cells were purified by FACS and GFP-positive cells were cultured in the presence of the AR ligand R1881 prior to harvesting and preparation of RNA. Expression of the androgen-regulated genes PSA and DRG-1 and a non-androgen-regulated gene GAPDH, was determined by PCR. (-) refers to uninfected cells treated with R1881. Arrows show positions of the expected size of the appropriate PCR product, as labelled on the left.

Figure 5. Schematic representation of the receptors and PLZF. The domain structure of retinoic acid receptor I (RARI) and oestrogen receptor I (ERI) are shown. Regions A-F refer to regions of differing amino-acid sequence homology as first described by Krust *et al* (1986) *EMBO J.* **5**, 891-897, the DNA binding domain (DBD, region C) and the ligand binding domain (LBD, region E) are most well conserved between the receptors. Transcription Activation Functions 1 and 2 (AF-1 and AF-2) refer to the regions of the receptor containing sequences required for transcriptional activation. The LBD also contains sequences which interact with co-repressor proteins. PLZF-RARI shows the fusion protein in acute promyelocytic leukaemia which results from the t(11;17) translocation, fusing the first 455 amino-acids of PLZF with regions B-F of RARI (described by Chen *et al* (1993) *EMBO J.* **12**, 1161-1167). The

PLZF-ER construct which we have made is shown, fusing the first 455 amino-acids of PLZF with ERI sequences (amino-acids 151-595) homologous to the RARI sequences present in PLZF-RARI. Also shown is HEG19 (see Tora *et al* (1989) *EMBO J.* 8, 1981-1986 which contains amino-acids 180-595 of human ERI. The open circles with the lettering Zn^{++} show the positions of the Kruppel-like $Zinc^{++}$-binding fingers present in PLZF.

Figure 6. Establishment of MCF-7 cell lines expressing PLZF-ER in a tetracycline-inducible manner. PLZF-ER was cloned into the tet-off plasmid pREVTRE (Clontech), which also confers resistance to hygromycin, and stably transfected into MCF7 tet-off (TO) cell line (Clontech). This line constitutively expresses the Tet repressor and will repress expression of genes pREVTRE, in the presence of doxytetracycline (Tet; 100 μg/ml) in the medium). MCF7 Pool is the pool of cells that survived hygromycin selection, whilst JP13 and JP23 are two representative and independent clones from that pool. For investigating the activity of an estrogen-regulated reporter gene, the parental MCF-7-TO, Pool, JP13 and JP23 cells were transiently transfected with a chloramphenicol acetyl transferase (CAT) reporter gene containing an ERE (ERE-G-CAT), following removal of Tet and/or the addition of 17ß-estradiol (E2; 10 nM). Panel (A) - Reporter gene activity was similar in MCF7-TO in the absence or presence of Tet and in the pool, JP13 and JP23 in the presence of Tet. In the absence of Tet, however, reporter gene activity was significantly reduced in the pool, JP13 and JP23. Results represent at least 3 independent experiments. In each experiment ER activity in the presence of E2 was taken as 100%. All other activities are shown relative to this. Panel (B) - Immunoblot of whole cell extracts prepared from the parental MCF-7-TO, Pool and clones JP13 and JP23

cells grown in the presence or absence of Tet using an anti-PLZF antibody.

Figure 7. PLZF-ER represses expression of the estrogen-responsive progesterone receptor gene. Total RNA was extracted from the parental MCF-7 cells (MCF-7 TO), MCF-7 Pool and the clones JP13 and JP23 and used as a template for an RT-PCR-based analysis of progesterone receptor (PR) mRNA levels. Each cell line was grown in the presence or absence of Tet (100 μg/ml), as shown. 17ß-estradiol (10 nM) was present throughout the course of the experiment. The expected RT-PCR product of PR was 238bp. The expected RT-PCR product of the house-keeping (control) gene GAPDH was 595bp. The upper panel shows RT-PCR data for PR and GAPDH, a housekeeping control gene mRNA levels, when cell lines are grown in the presence or absence of TET. Lower panel displays PR mRNA levels relative to GAPDH mRNA levels.

Figure 8. Expression of progesterone receptor expression in MCF7-Tet Off cells. The expression of progesterone receptor (PR) in the MCF-7 Tet Off cell line (A) and PLZF-ERα stably transduced MCF-7 Tet Off cell clones JP13 (B) and JP23 (C) was assessed by immunohistochemical staining. Prepared cells were immunostained with a mouse monoclonal antibody against human PR (Biogenics), with staining visualised using a secondary antibody (Vector laboratories) conjugated to Horseradish peroxidase (HRP). Positive staining is brown in colour and localised to the cell nucleus. Immunostained cells were counter stained with Haematoxylin - Eosin (H&E), which stains the cell nuclei blue and cytoplasm red.

Figure 9. Chromatin Immunopreciptation of Parental MCF-7 Tet-OFF cells and PLZF-ER stably transfected cell lines. Cell lines MCF7-TO or MCF7 JP13 were grown to 80% confluence in phenol red free DMEM, 5% DSS, P/S/G plus selection reagents and Dox as required. Cells were treated with E2 (10^{-8}M) or an ethanol control for 30 minutes at 37°C prior to formaldehyde cross-linking and sonication. All samples were immunoprecipitated with an anti-acetylated histone H4 antibody (Upstate Biotechnology, UK). DNA was recovered and subjected to PCR amplification. The expected size of the PCR product for the progesterone receptor (PR) gene was 260bp. Upper panel shows PR PCR product using DNA associated with immunoprecipitated chromatin (P) or total DNA extracted from the cells (S), using acetylated histone H4 antibody. Cells without (MCF7-TO) or with (MCF7 JP13) the PLZF-ER construct were grown in the presence (+) or absence (-) of TET. Lower panel displays PR PCR product expressed relative to the quantity of PR PCR product using DNA immunoprecipitated with an acetylated histone H4

Figure 10. Growth of MCF7-Tet Off line JP23 in response to PLZF-ER expression. MCF7 cell growth is dependent upon the presence of oestrogen in the growth media. Cell line MCF7-Tet Off JP23 expresses PLZF-ER in the absence of TET. Shown here is the growth rate of cell cultures of JP23 under various growth conditions. Oestradiol + TET D6 cells were grown without TET for 4 days then retreated with TET for 2 days.

Example 1: Construction and use of PLZF-AR gene fusions

Fig. 1 shows a schematic of the constructs used in the experiments. Also shown are results with reporter gene assays by transient transfection using these constructs. Of note is the fact that PLZF-AR represses in a ligand-

dependent manner. Swapping the ligand binding domain (LBD) for the estrogen receptor (ER) LBD but maintaining the androgen receptor DNA binding domain (DBD) to give PLZF-AR-ER (referred to as PANER) results in repression of androgen-regulated reporter genes in an estrogen or anti-estrogen dependent manner. Thus, inducible repressors may be made using AR and ER LBDs.

Fig. 2 shows western blots to confirm that proteins are made and are recognised by the appropriate antibodies.

Fig 3 shows transient transfection of the prostate cancer cell line LnCap with empty vector or PLZF-AR containing vector. This vector also encodes green fluorescent protein (GFP). Cell sorting using a FACS machine followed by cultuing and assaying for the classic prostate cancer marker PSA was then carried out. This demonstrates repression of an endogenous gene.

Figure 4 shows that PLZF-AR represses endogenous Androgen-regulated gene expression, using PCR to detect and quantify expression.

A 1392 bp region of PLZF coding region was amplified by PCR from a full length cDNA clone using a generic oligonucleotide primer to 5N flanking cloning vector sequence (T7 primer;) and a primer complementary to PLZF sequences encompassed by bases 1441–1446 of the sequence of Chen *et al* (1993) *EMBO J.* **12**, 1161-1167, with additional bases added to the 3N end, so as to introduce an in-frame XhoI restriction enzyme site (Primer PLZF R; ccgctcgagCTGAATGAGCCAGTAAGTGCATTCTC). The resultant cloned PLZF gene fragment encodes the first 455 amino acids of PLZF,

which can be fused in frame with, for example, the required steroid hormone receptor domain(s).

A 1146 bp region of a human AR cDNA clone (Tilley *et al* 1989 *Proc. Natl. Acad. Sci. USA* **86**, 327-331) was amplified by PCR using primers which introduced an in frame XhoI site into the 5N coding region and a BamHI site immediately following the stop codon (Primers AR F1; ggagctcgagggTTGGAGACTGCCAGGGACC and AR R1; gtgaggatccTCACTGGGTGTGGAAATAGATGG). The AR PCR product was restriction enzyme digested with XhoI and BamHI and ligated with XhoI/BamHI digested PLZF-ER to replace the ER portion with AR (see WO01/ 02019). The ligation product was used to transform *E. coli* bacteria and plasmid DNA prepared from individual clones. Recombinant pSG5 plasmids containing PLZF-AR gene fusion DNA were initially identified by restriction enzyme digestion and were subsequently confirmed by DNA sequence analysis. The resultant cloned PLZF-AR gene encodes the first 455 amino acids of PLZF, fused in frame with amino acids 537-917 of human AR. Transient transfections in COS-1 cells, followed by immunoblotting of cell extracts with antibodies directed against PLZF or AR were used to confirm expression and expected size.

Example 2: Patient treatment using a ligand-modulated molecule

A plasmid vector is produced which encodes an PLZF-AR fusion protein, for example under the control of the PSA gene promoter which allows for selective expression in prostate tissue. The plasmid is prepared in a sterile and pyrogen-free form and is formulated into liposomes. The plasmid DNA-containing liposomes are administered into the vicinity of the prostate. Plasmid DNA is taken up by the prostate cancer cells and

androgen receptor-mediated transcription is suppressed selectively in prostate cells. PLZF-AR activity is initiated by the administration of an androgen or the antiandrogen cyproterone acetate (CPA).

Suppression of oestrogen receptor-mediated transcription in a breast cancer patient

A retroviral vector is produced which encodes a PLZF-ER or PLZF-AR-ER fusion protein as described in Example 1. Following packaging the recombinant retrovirus is transduced into breast cacnere cells *in situ* and oestrogen receptor-mediated transcription is suppressed selectively in breast cells. PLZF-ER or PLZF-AR-ER expression can be under the control of a mammary-specific promoter such as MUC1. The retroviral vector is administered into the site of the breast tumour. Retroviral RNA is taken up by the breast cancer cells and oestrogen receptor-mediated transcription is suppressed selectively in breast cells. PLZF-AR-ER activity is initiated by the administration of an androgen or the antiandrogen cyproterone acetate (CPA).

Example 3: Establishment of MCF-7 cell lines expressing PLZF-ER in a tetracycline-inducible manner

In order to examine the effect of PLZF-ER upon the expression of endogenous estrogen-responsive genes, we first developed an inducible PLZF-ER gene expression system in a human breast cancer cell line and tested this on the level of activity of a reporter gene introduced into the same cell line. Experimental procedures described below are well known to those skilled in the art.

PLZF-ER (as discussed in Example 1 and illustrated in Figure 5) was cloned into the TET-off plasmid pREVTRE (Clontech), which also confers resistance to hygromycin, and the recombinant plasmid stably transfected into the MCF7 TET-off (TO) cell line (Clontech). MCF7 is a human breast cancer cell line. Growth of MCF7 is dependent upon the presence of oestrogen in the growth media.

In the presence of doxytetracycline (TET65; 100 μg/ml) in the growth medium, the transformed MCF7 cell lines constitutively expresses the TET repressor and represses expression of genes held in the pREVTRE plasmid. Three different populations of cells were selected for further investigation. MCF7 Pool is a pool of cells that survived hygromycin selection, whilst JP13 and JP23 are two representative and independent clones from that pool.

MCF-7-TO (the parental cell line), MCF7 Pool, JP13 and JP23 cells were transiently transfected with an estrogen-regulated reporter gene, chloramphenicol acetyl transferase (CAT), containing an ERE (ERE-G-CAT). Following removal of TET and/or the addition of 17ß-estradiol (E2; 10 nM) reporter gene activity was assessed and the data is presented in Figure 6. Reporter gene activity was similar in MCF7-TO in the absence or presence of TET and in the MCF7 pool, JP13 and JP23 cell lines in the presence of TET. In the absence of TET, however, reporter gene activity was significantly reduced in the MCF7 pool, JP13 and JP23 cell lines. The results represent at least 3 independent experiments. In each experiment ER activity in the presence of E2 was taken as 100%. All other activities are shown relative to this.

To confirm that the changes in CAT activity were associated with PLZF-ER protein levels, immunoblotting of whole cell extracts prepared from MCF-7-TO, MCF7 pool, JP13 and JP23 cells grown in the presence or absence of TET was performed using an anti-PLZF antibody. The results are presented in Figure 6, panel B.

The data shows that TET+ oestrogen (E2)+ cells have CAT reporter gene activity. This is due to the TET repressing the expression of PLZF-ER from the pREVTRE plasmid and E2 inducing transcription of the oestrogen-responsive CAT reporter gene.

When TET is omitted from the growth media then PLZF-ER is transcribed and, in the case of E2+ cells, the PLZF-ER protein is activated and blocks transcription of the CAT reporter gene. In the absence of E2 no CAT activity is recorded as oestrogen is required to activate the reporter gene and to activate PLZF-ER.

Hence MCF-7 cell lines JP13 and JP23 express PLZF-ER in a tetracycline inducible manner. Also, expression of PLZF-ER is associated with a decrease in CAT reporter gene activity in an E2 – dependent manner.

Example 4: Expression of an endogenous oestrogen-responsive gene in PLZF-ER cell lines

In Example 3 we established that the tetracycline-inducible PLZF-ER construct could down-regulate the activity of an oestrogen-responsive reporter gene. We now describe a series of experiments that demonstrate that the PFZF-EF construct can regulate the expression of progesterone receptor (PR), an endogenous oestrogen – responsive gene. The gene

expression levels were analysed using RT-PCT, a technique well known to those skilled in the art.

MCF-7 cells (MCF-7 TO), MCF-7 Pool cell and the cell lines JP13 and JP23 (described in Example 3) were grown in the presence or absence of TET (100 μg/ml), with 17ß-estradiol (10 nM) present throughout the course of the experiment. Total RNA was extracted from the cells using Qiagen RNAeasy kits (Qiagen, UK). 2 μg of total RNA was used to make randomly primed cDNA. This cDNA was then used as a template for a PCR reaction using primers for progesterone receptor (PR) and GAPDH. The PR and GAPDH primers used were:

		Accession No.	bp
PR primers:			
PR-F2	AAATCATTGCCAGGTTTTCG	AF016381	2375-2395
PR-R2	TGCCACATGGTAAGGCATAA	AF016381	2584-2564
GAPDH Primers:			
GAPDH-F	CCACCCATGGCAAATTCCATGGCA	NM002046	229-252
GAPDH-R	TCTAGACGGCAGGTCAGGTCCAC	NM002046	802-824

The PCR conditions used were: 27 cycles, annealing temp. 50°C for the PR RT-PCR product of 238bp; 30 cycles, annealing temp. 55°C for the house-keeping (control) gene GAPDH RT-PCR product of 595bp.

The data from this experiment is shown in Figure 7. It can be clearly seen that when TET is omitted from the cell growth conditions (-) there is a reduction in the expression of the PR gene. Hence in cell lines where PLZF-ER is expressed (see Figure 6 panel B for PLZF-ER expression in TET- cell lines) there is a reduction in the expression of an endogenous oestrogen – responsive gene.

In order to established whether there was a reduction in progesterone receptor (PR) protein levels in TET- cells, we conducted an immunohistological staining for PR protein in these cells, a technique well known to those skilled in the art.

MCF-7 Tet Off cell line and PLZF-ERα stably transduced MCF-7 cell lines JP13 and JP23 (described in Example 3) were grown in complete medium in the presence of 10^{-8} M E2 and Doxytetracycline (1μg /ml). The cells were harvested by trypsinisation and subsequently spun onto the surface of glass microscope slides ("cytospun"), air dried, fixed and immunostained with a mouse monoclonal antibody against human PR (Biogenics), with staining visualised using a secondary antibody (Vector laboratories) conjugated to Horseradish peroxidase (HRP). Positive staining is brown in colour and localised to the cell nucleus. Immunostained cells were counter stained with Haematoxylin - Eosin (H&E), which stains the cell nuclei blue and cytoplasm red. The results are shown in Figure 8. Cells are shown at a magnification of x40.

As seen in panel (A), strong nuclear expression of PR is seen in MCF7-Tet Off cells. Also detectable in these cells is weaker, cytoplasmic staining, which is known to be a background staining effect. In comparison, nuclear expression of PR is greatly reduced in the PLZF-ERα stable lines JP13 (B) and JP23 (C). These lines, however, show some weak cytoplasmic, background, staining seen in the parental line.

Hence expression of the PLZF-ER construct is associated with a reduction in the protein levels of PR, an endogenous oestrogen-responsive gene.

Example 5: PLZF-ER mediates a change in the histone acetylation state of DNA associated with the progesterone receptor gene

As seen in Examples 3 and 4, the PLZF-ER protein can reduce the level of gene expression of both an introduced oestrogen responsive reporter gene (CAT) and an endogenous oestrogen responsive gene (PR), leading to the reduction in nuclear protein levels of the latter. We propose that this is due to the PLZF- portion of the polypeptide mediating a change in the histone acetylation state of DNA at or close to where the ER-portion of the polypeptide binds. To demonstrate this, we decided to assess the quantity of acetylated histone H4-assocated PR DNA in cell lines in which the PLZF-ER construct was assessed using chromatin immunoprecipitation, a technique well known to those skilled in the art.

Chromatin immunoprecipitation (ChIP) analysis was carried out using a ChIP Assay kit according to the manufacturer's instructions (Upstate Biotechnology, UK). MCF-7 Tet Off cell line and PLZF-ERα stably transduced MCF-7 cell line JP13 (described in Example 3) were grown to 80% confluence in phenol red free DMEM, 5% DSS, P/S/G plus selection reagents and Dox as required. Cells were treated with E2 (10^{-8}M) or an ethanol control for 30 minutes at 37°C prior to formaldehyde cross-linking and sonication. All samples were immunoprecipitated with an anti-acetylated histone H4 antibody (Upstate Biotechnology, UK). DNA was recovered and the quantity of progesterone receptor (PR) DNA assessed by PCR amplification using the primers shown below:

		Accession No.	bp
PR SA1465	AAAGGGGAGTCCAGTCGTCATG	X51730	1436-1457
PR SA1466	TGCTGGTCCTGCGTCTTTTC	X51730	1687-1668

The PCR condition used was 35 cycles, annealing temp. 54.5°C for the progesterone receptor (PR) gene product of 260bp.

The data from this experiment is shown in Figure 9. The products from the PCR reaction are shown in the upper panel. P indicates the DNA sample used as PCR template was derived from anti-acetylated histone H4 antibody immunoprecipitated DNA, whilst S indicates the DNA sample used as PCR template was derived from total DNA extracted from the cells. (-) refers to DNA recovered from immunoprecipitations where the anti-acetylated histone H4 antibody was omitted. The lower bar chart shows levels of PCR product obtained relative to the immunoprecipitated DNA for the parental cells (MCF7-TO).

From this data it can be seen that PLZF-ER polypeptide reduces the quantity of PR PCR product in immunoprecipitated DNA (compare the PCR product of JP13 cell line sample P +TET to the quantity of PCR product in JP13 cell line sample P -TET). It is likely that the reduction in the quantity of PCR product derived from immuoprecipitated DNA extracted from JP13 cells grown without TET is due to the PLZF portion of the polypeptide mediating a deacetylation of histone H4 proteins at or close to where the ER-portion of the polypeptide binds. Less PR gene DNA is immunoprecipitated as the antibody used for the immunoprecipitation can only recognise acetylated histone H4. From this we propose that the reduction in PR gene transcript (see Example 4 and Figure 7) and protein levels (see Example 4 and Figure 8) in cells that express PLZF-ER and grown in the presence of oestrogen (E2) is caused by the PLZF-ER protein binding to the oestrogen-response element in PR and recruiting a histone deacetylation complex (HDAC) to the gene which results in PR repression due to histone deacetylation.

Example 6: Affect of PLZF-ER on the growth rate of MCF-7 cell line JP13

The MCF7-Tet Off line JP23 expresses PLZF-ER in the absence of tetracycline (see Example 3 and Figure 6, panel B). MCF7 is a human breast cancer cell line whose growth is dependent upon the presence of oestrogen in the growth media.

We wished to evaluate the effect of PLZF-ER on oestrogen regulated growth responses in this line. MCF7-Tet Off line JP23 cells were seeded in an oestrogen free media containing Doxytetracycline (TET 1μg / ml) and maintained for two days. Cell culture media were then altered so that oestrogen-regulated growth could be assessed in the presence (no TET) and absence (plus TET) of PLZF-ER. As shown in Figure 10, oestrogen regulated growth is inhibited in JP23 cells in the presence of PLZF-ER expression (oestradiol only versus oestradiol + TET). In order to address if this growth inhibition was permanent, cells expressing PLZF-ER for four days were re-treated with TET. Under these conditions PLZF-ER expression is lost within 24 hours. However, over the remaining 5 days of the assay the cells showed little growth, indicating a greatly reduced growth response to oestrogen.

Hence the PLZF-ER polypeptide can block the oestrogen-dependent growth of MCF7 cells, and that this repression is partially reversible. We propose that the PLZF-ER polypeptide is repressing the expression of genes that are responsible for the oestrogen-dependent growth of MCF7

cells in the same way as the polypeptide repressed the progesterone receptor gene (see Example 5).

CLAIMS

1. A method of suppressing the expression of a selected endogenous gene in a eukaryotic cell, the method comprising introducing into the cell (a) a molecule comprising a nucleic acid binding portion which binds to a site at or associated with the selected endogenous gene and a modifying portion which comprises a polypeptide or peptidomimetic which is capable of modulating covalent modification of nucleic acid or chromatin, or (b) a polynucleotide encoding said molecule.

2. The method of claim 2 wherein the modifying portion is a chromatin inactivation portion.

3. The method of claim 1 wherein the modifying portion is capable of modulating methylation of DNA.

4. The method of claim 3 wherein the modifying portion comprises a methyl transferase or a component of a methyltransferase complex,

5. A method according to any of the preceding claims wherein the nucleic acid binding portion is a DNA binding portion.

6. A method according to Claim 1 wherein the nucleic acid binding portion is an RNA binding portion and the site present in a eukaryotic genome is a nascent RNA being transcribed from DNA.

7. A method according to any one of Claims 2, 5 to 6 wherein the chromatin inactivation portion facilitates histone deacetylation.

8. A method according to any one of Claims 2, 5 to 7 wherein the chromatin inactivation portion is all or a portion of a component of a histone deacetylation (HDAC) complex or all or a portion of a polypeptide which binds to or facilitates the recruitment of a HDAC complex.

9. A method according to Claim 8 wherein the component of the HDAC complex or the polypeptide which binds to or facilitates the recruitment of a HDAC complex is any one of PLZF, N-CoR, SMRT, Sin3, SAP18, SAP30 and HDAC.

10. A method according to Claim 9 wherein the chromatin inactivation portion is all or a N-CoR- or SMRT-binding part of PLZF.

11. A method according to Claim 9 wherein the chromatin inactivation portion is all or an enzymatically active part of a HDAC.

12. A method according to any one of Claims 1 to 5, 7 to 11 wherein the DNA binding portion is all or a DNA-binding part of a zinc-finger DNA binding protein or all or a DNA-binding part of a helix-turn-helix DNA binding protein.

13. A method according to Claim 12 wherein the DNA binding portion is all or a DNA-binding part of an animal or plant DNA binding protein.

14. A method according to Claim 5, 7 to 12 wherein the DNA binding portion is all or a DNA-binding part of a bacterial or yeast DNA binding protein engineered to bind plant or animal genome.

15. A method according to any one of Claims 5, 7 to 10 wherein the DNA binding portion is all or a DNA binding part of a steroid hormone receptor protein or other nuclear receptor DNA binding protein.

16. A method according to Claim 15 wherein the DNA binding portion is all or a DNA-binding portion of estrogen receptor (ER) or all or a DNA-binding portion of androgen receptor (AR).

17. A method according to any one of the preceding claims wherein the binding of the nucleic acid binding portion to the target site is capable of being regulated by a ligand.

18. The method of any of the preceding claims wherein the molecule comprises a ligand binding portion.

19. A method according to claim 17 or 18 the method further comprising the step of exposing the cell to the ligand.

20. The method of any one of claims 17 to 19, wherein the molecule does not comprise both the ligand binding domain and the DNA binding domain of the androgen receptor (AR).

21. The method of any one of claims 17 to 20 wherein the nucleic acid binding portion and ligand binding portion are derivable from different polypeptides, for example from different steroid hormone receptors.

22. A molecule, for example polypeptide, comprising a nucleic acid binding portion which binds to a site at or associated with a eukaryotic gene, a ligand binding portion and a modulating portion which comprises

a polypeptide or peptidomimetic which is capable of modulating covalent modification of nucleic acid or chromatin, wherein the ability of the molecule to bind to the site is capable of being regulated by binding of a ligand to the ligand binding portion, wherein the nucleic acid binding portion and ligand binding portion are not derivable from the same naturally occuring molecule.

23. A polynucleotide encoding a polypeptide according to claim 22.

24. A method of suppressing the expression of a selected gene in a eukaryotic cell, the method comprising introducing into the cell (a) a molecule comprising a nucleic acid binding portion which binds to a site at or associated with the selected gene which site is present in a eukaryotic genome and a modifying portion which comprises a polypeptide or peptidomimetic which is capable of modulating covalent modification of nucleic acid or chromatin, for example a chromatin inactivation portion, or (b) a polynucleotide encoding said molecule, wherein the ability of the molecule to bind to the site is capable of being regulated by a ligand, preferably a small molecule, and the method further comprises the step of exposing the cell to the ligand, wherein the molecule does not comprise both the ligand binding domain and the DNA binding domain of the androgen receptor (AR).

25. The method of claim 24 wherein the molecule comprises a ligand binding portion.

26. The method of claim 18 or 25 or molecule of claim 22 or polynucleotide of claim 23 wherein the ligand binding portion comprises a ligand binding portion of the androgen receptor (AR).

27. The method or molecule or polynucleotide of any one of claims 22 to 26 wherein the nucleic acid binding portion comprises a DNA binding portion of the estrogen receptor (ER).

28. A method according to Claim 6 wherein the RNA binding protein binds to nascent RNA expressed from proviral DNA.

29. A method according to Claim 28 wherein the RNA binding protein is *tat* or a *tat*-like protein or an RNA-binding portion thereof.

30. A method or molecule or polynucleotide according to any one of the preceding claims wherein the nucleic acid binding portion and the modifying portion or chromatin inactivation portion are fused.

31. A molecule or polypeptide or polynucleotide according to any one of claims 22 to 27, 30 wherein the nucleic acid binding portion or the modifying or chromatin inactivation portion are as defined in any one of Claims 1 to 21.

32. A method or polynucleotide according to any one of claims 1 to 21, 23 to 31 wherein the polynucleotide comprises a promoter operably linked to allow expression of the polypeptide.

33. A method or polynucleotide according to Claim 32 wherein the promoter is an inducible promoter.

34. A vector comprising a polynucleotide according to any one of claims 23, 26 to 27, 30 to 33.

35. A method according to any of claims 1 to 21 or 24 to 30, 32, 33 wherein the polynucleotide is a vector comprising a polynucleotide as defined in any of those claims.

36. A method according to Claim 35 or vector according to claim 34 wherein the vector is an animal cell vector.

37. A method according to claim 35 or vector according to claim 34 wherein the vector is a plant cell vector.

38. A method according to any one of Claims 35 to 37 or vector according to claim 34, 36 or 37 wherein the vector is a viral vector.

39. A method according to any one of Claims 35 to 37 or vector according to claim 34, 36 or 37 wherein the vector is a plasmid vector.

40. A host cell comprising a polynucleotide according to any one of Claims 23, 26 to 27, 30 to 33 or a vector according to any one of Claims 34 or 36 to 39.

41. A host cell according to Claim 40 which is a bacterial cell.

42. A host cell according to Claim 40 which is an animal cell.

43. A host cell according to Claim 40 which is a plant cell.

44. An animal comprising a host cell according to Claim 42.

45. A plant comprising a host cell according to Claim 43.

46. A method according to any of claims 1 to 21, 24 to 30, 32, 33, 35 to 39 wherein the eukaryotic cell is an animal cell and is contained within an animal or is a plant cell and is contained within a plant.

47. A method according to any one claims 1 to 21, 24 to 30, 32, 33, 35, 36, 38, 39, 46 wherein the expression of a selected gene in a human is suppressed.

48. A method according to any one of claims 1 to 21, 24 to 30, 32, 33, 35 to 39, 46 wherein the expression of a plurality of selected genes is suppressed.

49. Use of a molecule or vector as defined in any one of the preceding claims in the manufacture of an agent for suppressing the expression of the selected endogenous gene in a eukaryotic cell.

50. Use of a molecule or polynucleotide or vector according to any one of claims 22, 26, 27, 30 to 34, 36 to 39 in the manufacture of an agent for suppressing the expression of the selected gene in a eukaryotic cell.

51. Use of a polynucleotide encoding a molecule as defined in any one of the preceding claims in the manufacture of an agent for suppressing the expression of the selected endogenous gene in a eukaryotic cell.

52. Use according to any one of claims 49 to 51 wherein the agent is a medicament for suppressing the expression of a selected gene in an animal.

53. A method of treating a patient in need of suppression of the expression of a selected gene, the method comprising administering to the patient an effective amount of a molecule or polynucleotide or vector as defined in any one of claims 49 to 51.

54. Use of a molecule or polynucleotide or vector as defined in any one of claims 49 to 51 in the manufacture of a medicament for suppressing the expression of a selected gene in a patient in need of such suppression.

55. The method of claim 53 or use of claim 54 wherein the molecule or polynucleotide or vector is as defined in claim 50.

56. A molecule or polynucleotide or vector as defined in claim 50 for use in medicine.

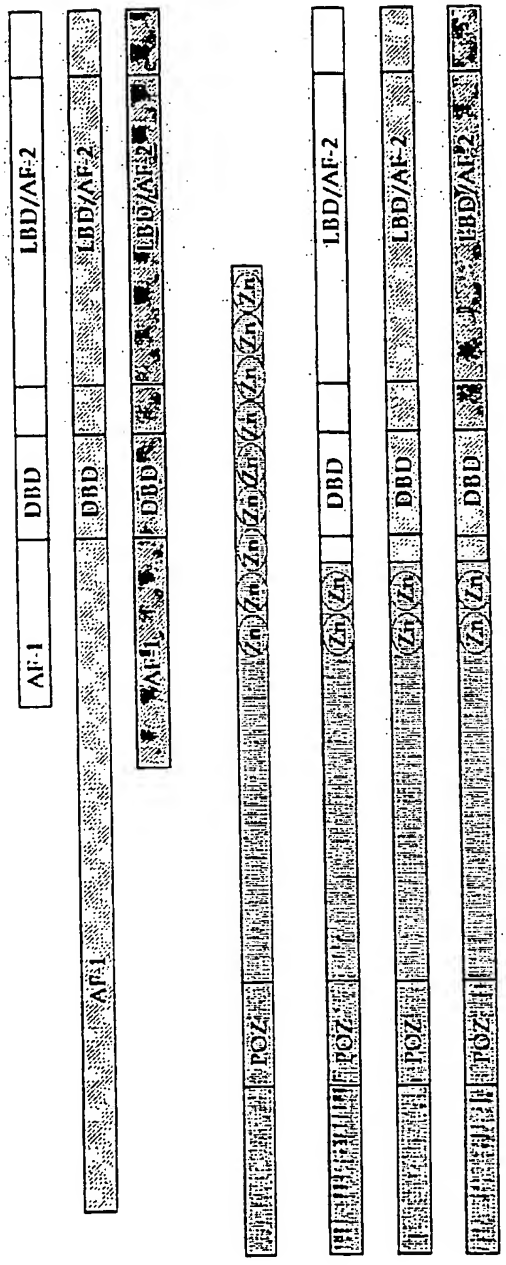
57. A pharmaceutical composition comprising a molecule or polynucleotide or vector as defined in claim 50 and a pharmaceutically acceptable carrier.

58. Any novel method of suppressing the activity of a selected gene in a plant or animal cell as herein described.

59. Any novel polypeptide which suppresses the activity of a selected gene in a plant or animal cell as herein described.

60. Any novel polynucleotide encoding a polypeptide which suppresses the activity of a selected gene in a plant or animal cell as herein described.

Figure 1: page 1



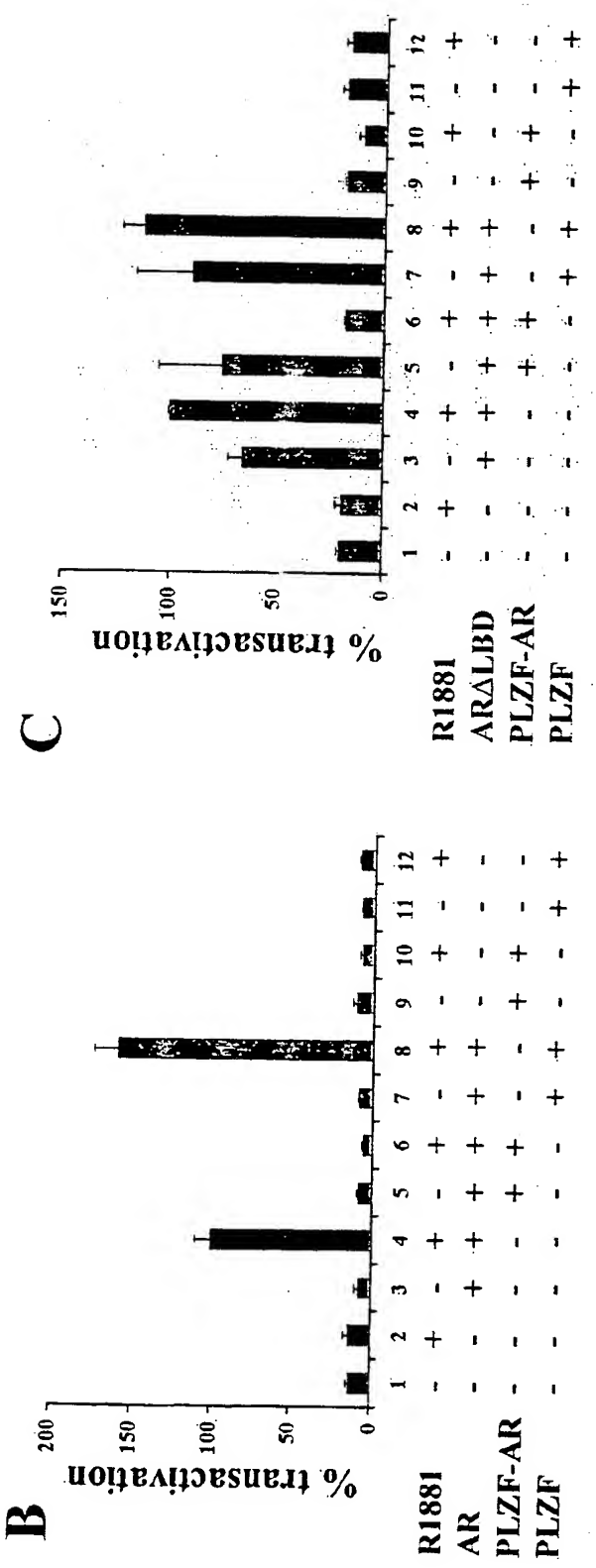


Figure 1: page 2

Figure 1: page 3

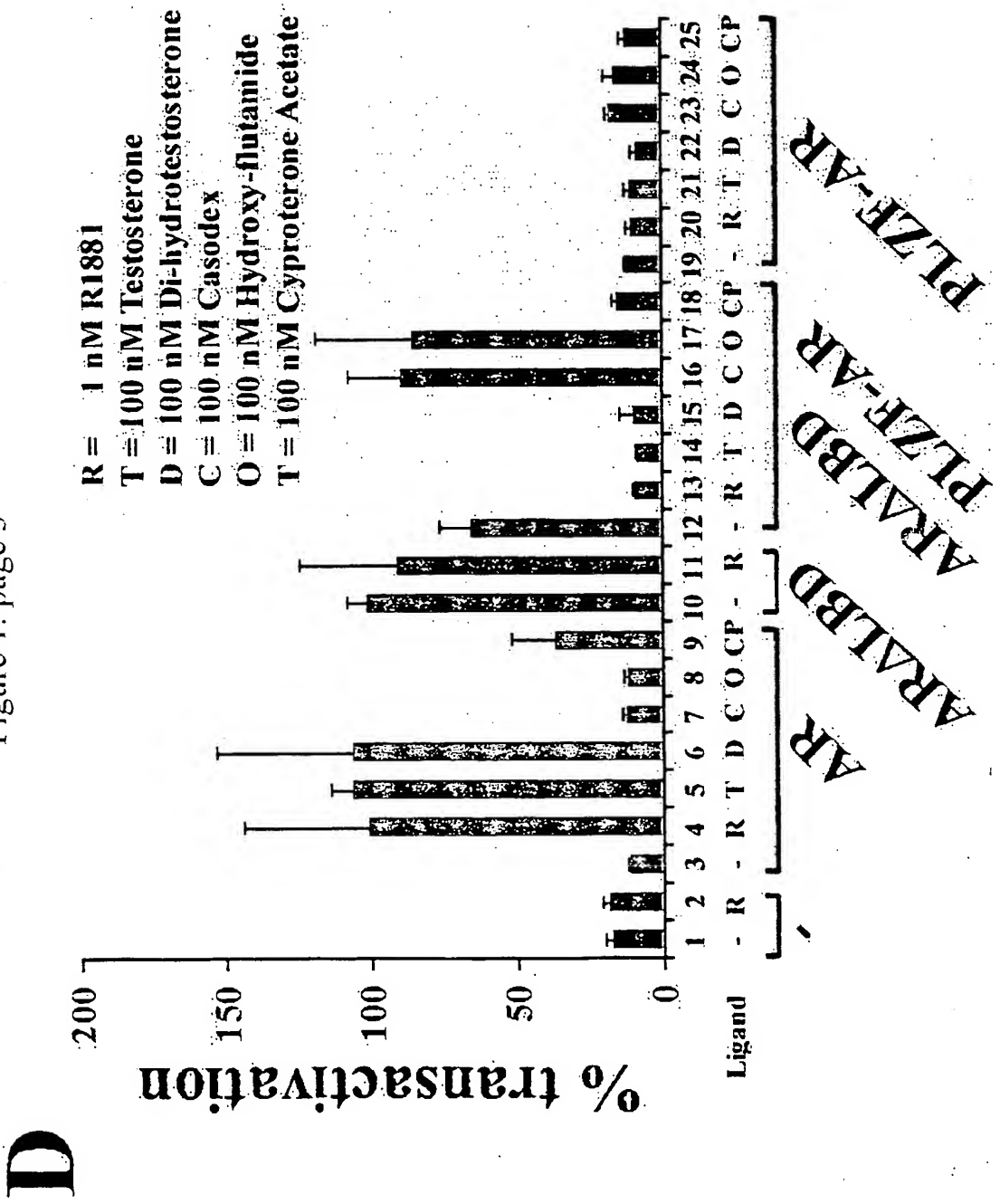


Figure 1: page 4

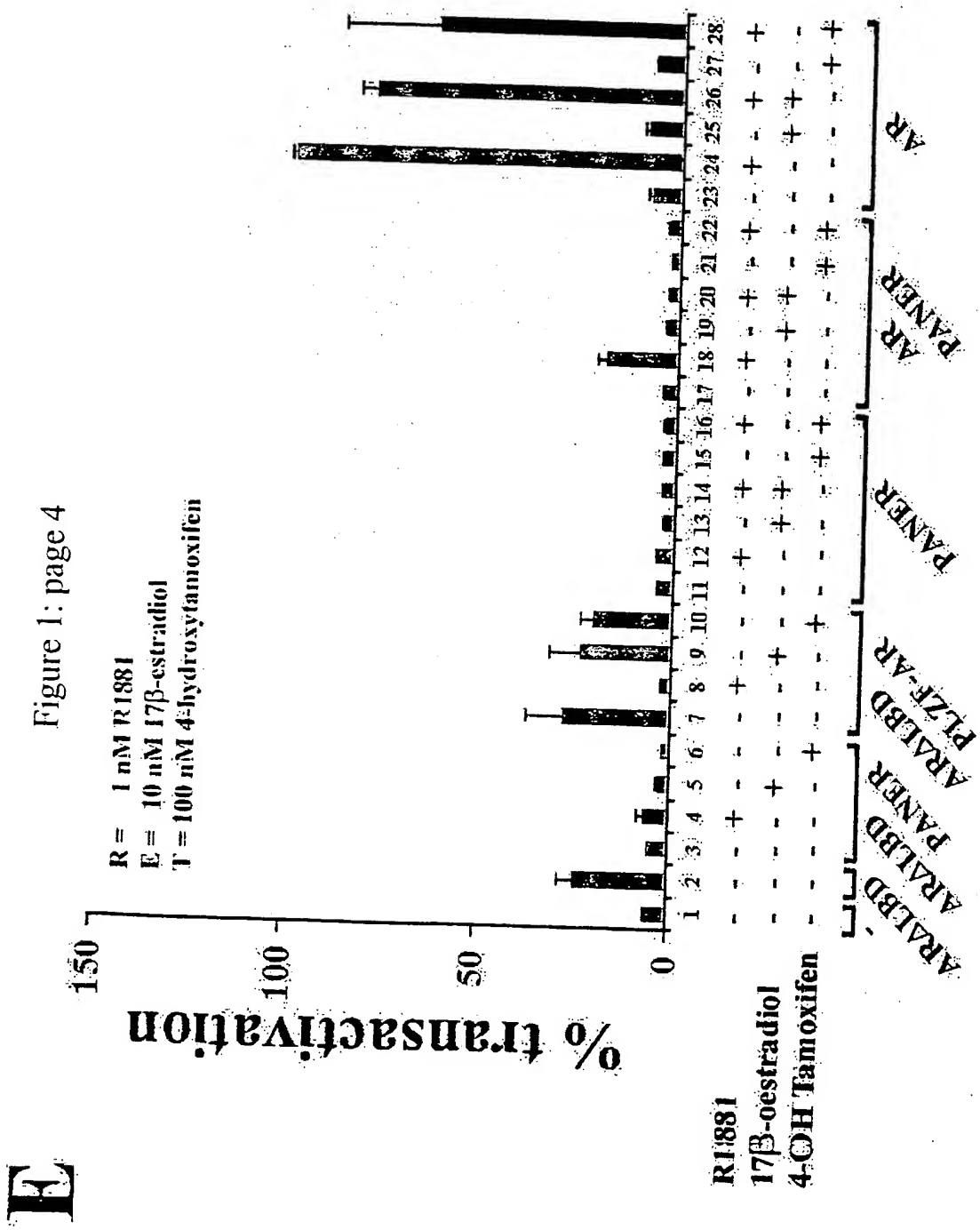


Figure 2

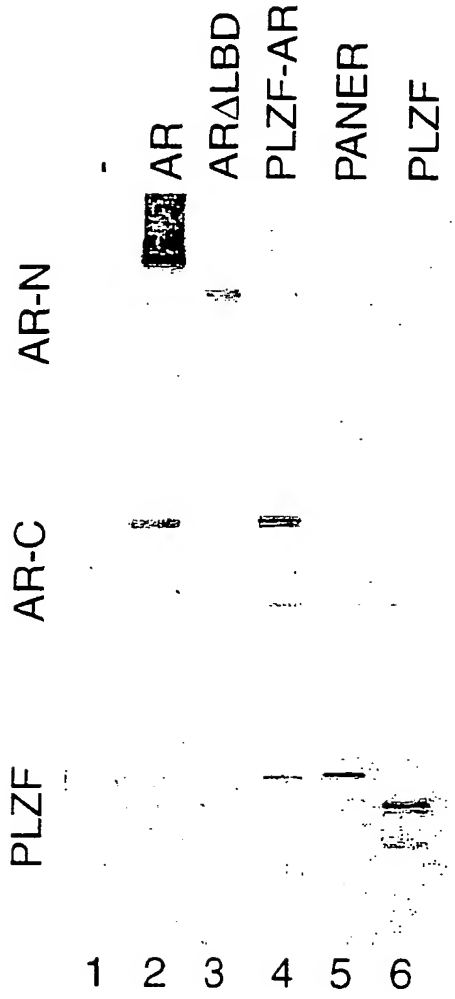


Figure 3

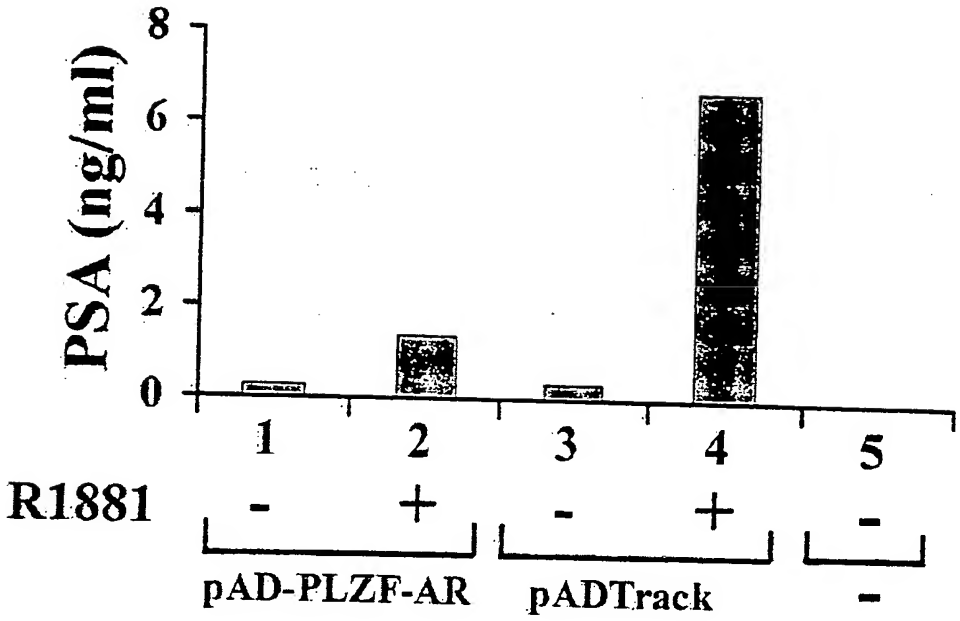


Figure 4

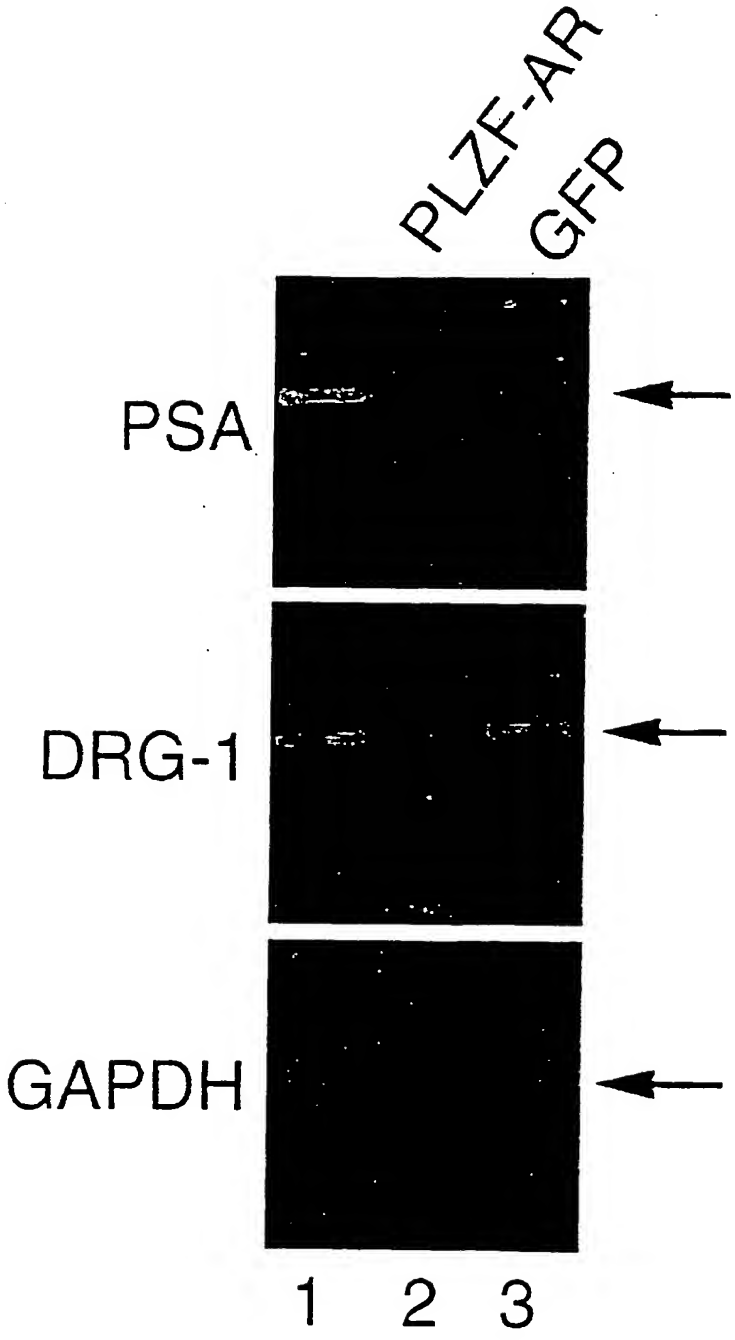


Figure 5

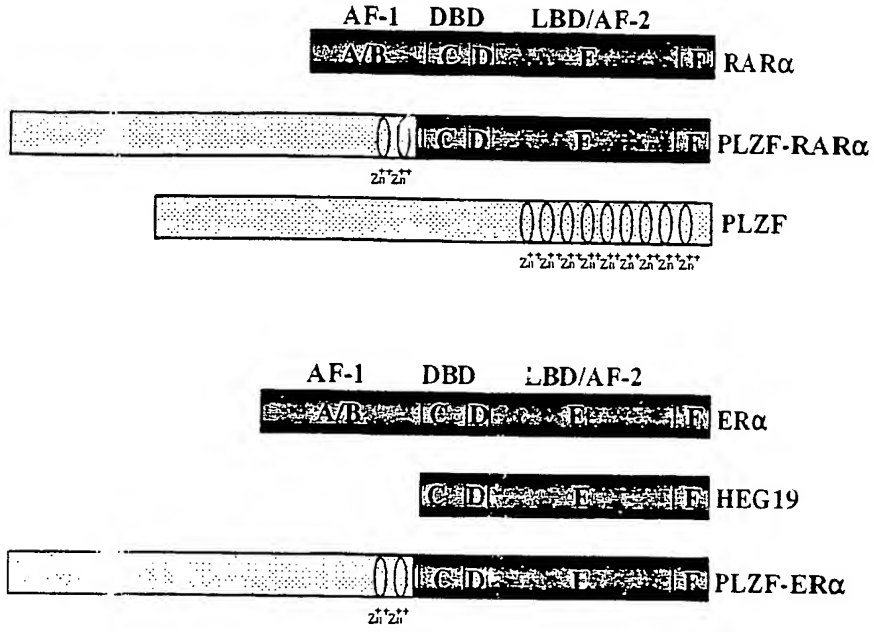


Figure 6

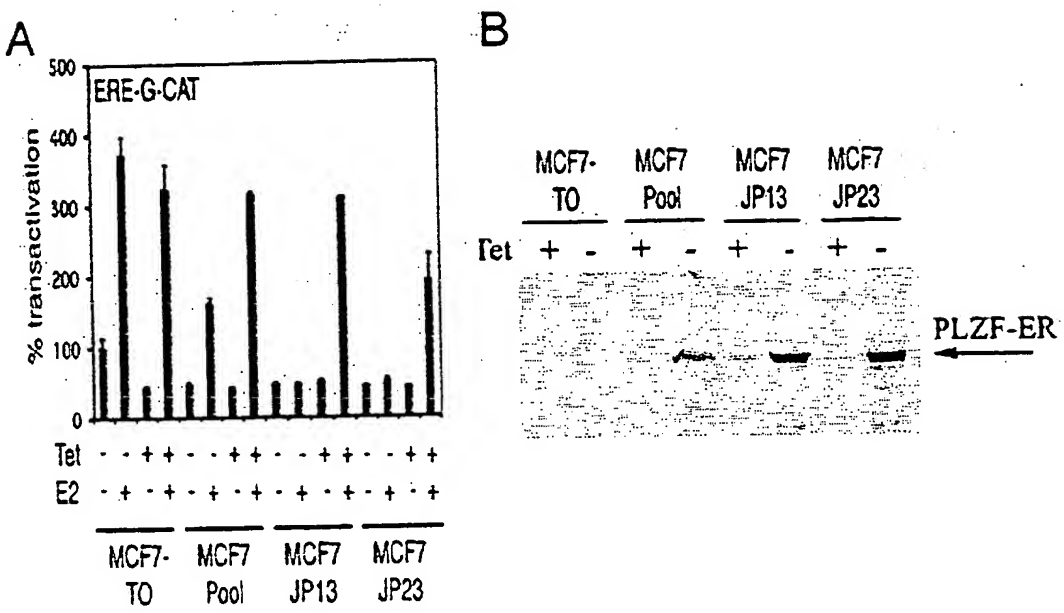


Figure 7

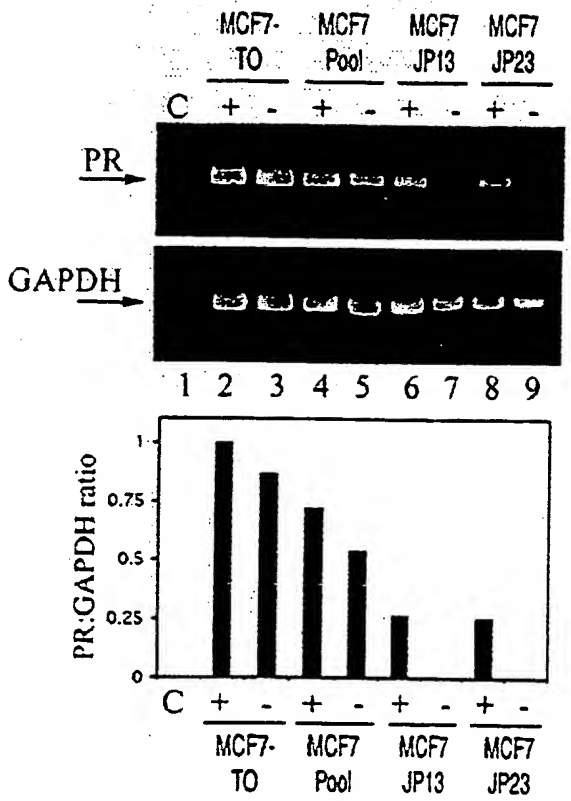


Figure 8

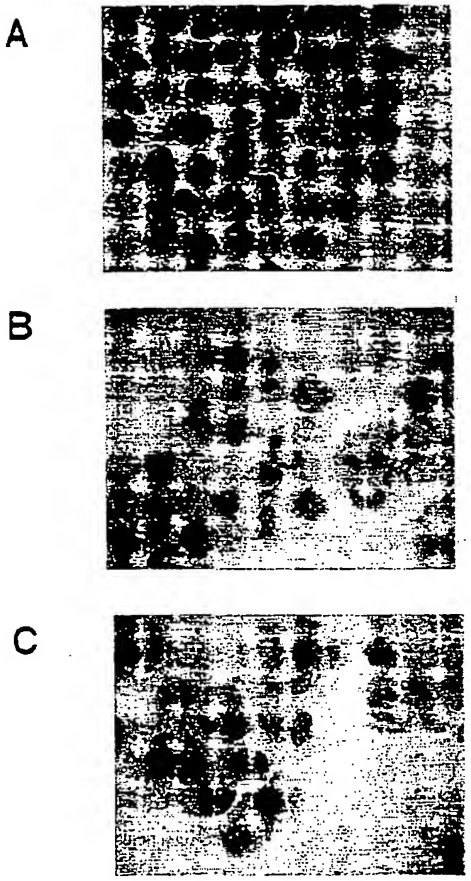


Figure 9

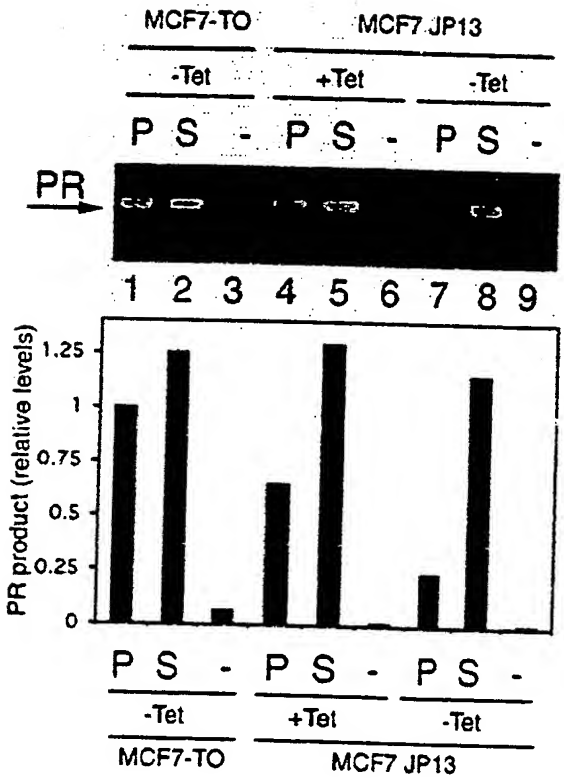
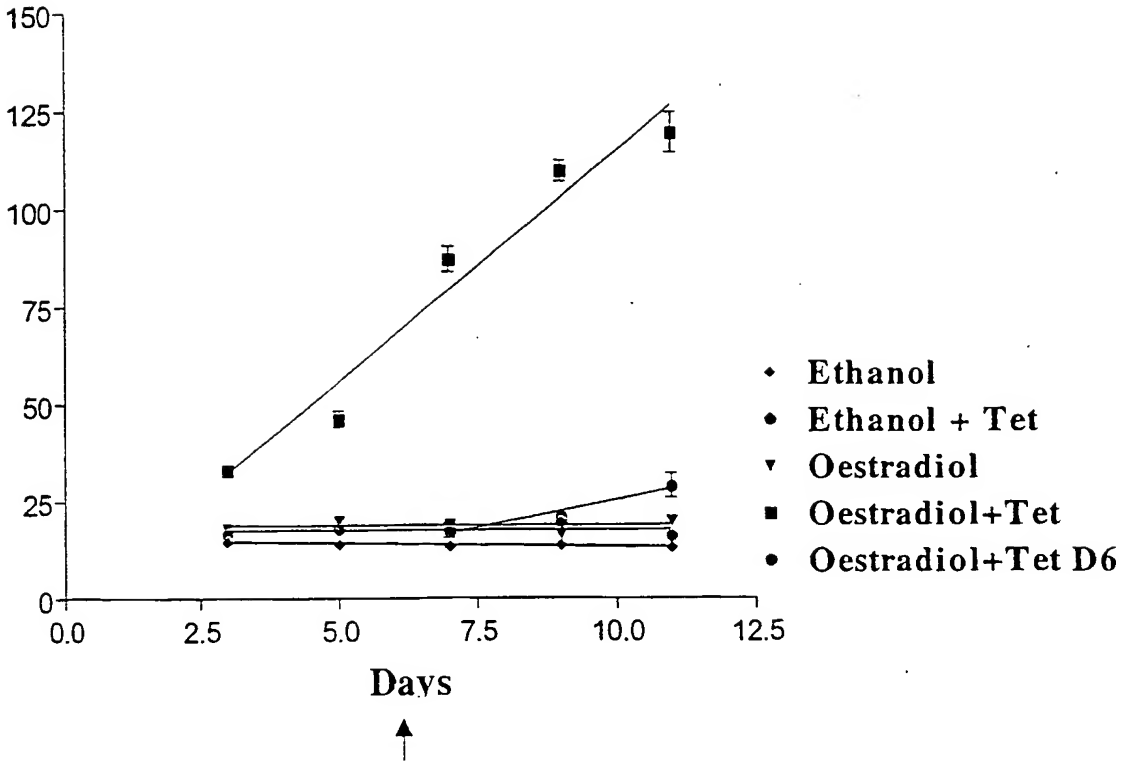


Figure 10


150
125
100
75
50
25
0
0.0
2.5
5.0
7.5
10.0
12.5
Days
Ethanol
Ethanol + Tet
Oestradiol
Oestradiol+Tet
Oestradiol+Tet D6
Oestradiol + Tet added on Day 6

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 February 2003 (06.02.2003)

PCT

(10) International Publication Number
WO 03/010308 A3

(51) International Patent Classification[7]: C12N 15/11, C07K 14/47, C12N 15/67, A61K 38/16, 47/48, C12N 5/10, A01K 67/027, C12N 15/82

(21) International Application Number: PCT/GB02/03336

(22) International Filing Date: 19 July 2002 (19.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0117964.7 24 July 2001 (24.07.2001) GB

(71) Applicant (*for all designated States except US*): GENE EXPRESSION TECHNOLOGIES LIMITED [GB/GB]; 90 Fetter Lane, London EC4A 1JP (GB).

(72) Inventors; and

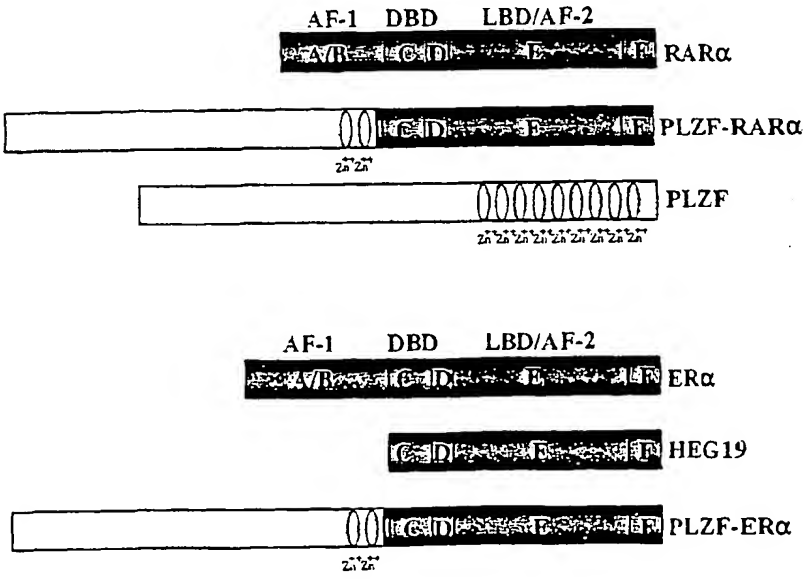
(75) Inventors/Applicants (*for US only*): BULUWELA, Lakjaya [GB/GB]; Imperial College School of Medicine, Dept. of Cancer Medicine, Du Cane Road, London W12 0NN (GB). HOLMES, David [GB/GB]; Imperial College School of Medicine, Dept. of Cancer Medicine, Du Cane Road, London W12 0NN (GB). KAMALATI, Tahereh [GB/GB]; Imperial College School of Medicine Dept. of Cancer Medicine, Du Cane Road, London W12 0NN (GB). WAXMAN, Jonathan [GB/GB]; Imperial College School of Medicine, Dept. of Cancer Medicine, Du Cane Road, London W12 0NN (GB). ALI, Simak [GB/GB]; Imperial College School of Medicine, Dept. of Cancer Medicine, Du Cane Road, London W12 0NN (GB).

(74) Agent: MILES, John, S.; Eric Potter Clarkson, Park View House, 58 The Ropewalk, Nottingham NG1 5DD (GB).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

[Continued on next page]

(54) Title: CONTROL OF GENE EXPRESSION



(57) Abstract: A method of suppressing the expression of a selected endogenous gene in a eukaryotic cell the method comprising introducing into the cell (a) a polypeptide comprising a nucleic acid binding portion which binds to a site at or associated with the selected gene and a modifying portion which comprises a polypeptide or peptidomimetic which is capable of modulating covalent modification of nucleic acid or chromatin, for example a chromatin inactivation portion; or (b) a polynucleotide encoding said polypeptide. The binding of the molecule to nucleic acid may be modulated by a ligand and the molecule may comprise a ligand binding portion. The nucleic acid binding portion may be a nucleic acid binding portion of a nuclear receptor DNA binding protein.

WO 03/010308 A3

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) **Date of publication of the international search report:**
8 May 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

Inte tal Application No
PCT/GB 02/03336

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/11 C07K14/47 C12N15/67 A61K38/16 A61K47/48
C12N5/10 A01K67/027 C12N15/82

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K A61K A01K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CHIEN P-Y ET AL: "A FUSION PROTEIN OF THE ESTROGEN RECEPTOR (ER) AND NUCLEAR RECEPTORCOREPRESSOR (NCOR) STRONGLY INHIBITS ESTROGEN-DEPENDENT RESPONSES IN BREAST CANCER CELLS" MOLECULAR ENDOCRINOLOGY, BALTIMORE, MD, US. vol. 13, no. 12, December 1999 (1999-12), pages 2122-2136, XP000978743 ISSN: 0888-8809 the whole document ---	1-57
X	WO 01 02019 A (IMP COLLEGE INNOVATIONS LTD ;ALI SIMAK (GB); BULUWELA LAKJAYA (GB)) 11 January 2001 (2001-01-11) the whole document ---	1-16, 28-32, 35-39, 58-60

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 March 2003	28/03/2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk: Te., (+31-70) 340-204C. Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Seranski, P

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inte: nal Application No
PC I/uB 02/03336

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	MANTEUFFEL-CYMBOROWSKA MALGORZATA: "Nuclear receptors, their coactivators and modulation of transcription." ACTA BIOCHIMICA POLONICA, vol. 46, no. 1, 1999, pages 77-89, XP009005858 ISSN: 0001-527X page 83 -page 84 ---	17-27, 32-57
A	RAZIN AHARON: "CpG methylation, chromatin structure and gene silencing - a three-way connection." EMBO (EUROPEAN MOLECULAR BIOLOGY ORGANIZATION) JOURNAL, vol. 17, no. 17, pages 4905-4908, XP002233552 ISSN: 0261-4189 the whole document -----	1-16, 28-32, 35-39, 58-60

INTERNATIONAL SEARCH REPORT

national application No.
PCT/GB 02/03336

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: –
because they relate to subject matter not required to be searched by this Authority, namely:

see FURTHER INFORMATION sheet PCT/ISA/210

2. [X] Claims Nos.: 58-60
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 1-21, 24-30, 32, 33, 35-39, 46-48, 53-55, 56 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.1

Rule 39.1(iv) PCT - Method for treatment of the human or animal body by therapy

---

Continuation of Box I.2

Claims Nos.: 58-60

Present claims 58-60 relate to an extremely large number of possible products and methods. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the compounds and methods claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely those parts relating to the products and methods methods of claims 1-57.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

ormation on patent family members

Int onal Application No
PCT/GB 02/03336

Patent documen: cited in search report	Publication date	Patent family member(s)	Publication date
WO 0102019 A	11-01-2001	AU 5556600 A	22-01-2001
		EP 1190073 A2	27-03-2002
		WO 0102019 A2	11-01-2001
		GB 2367555 A	10-04-2002
		JP 2003503082 T	28-01-2003

Form PCT/ISA/210 (patent family annex) (July 1992)